

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	120	100.0	749	2	Q72391	homo sapien
2	120	100.0	1103	2	Q6MZP4	homo sapien
3	120	100.0	2193	2	Q6NZM7	homo sapien
4	120	100.0	2240	2	Q6SDP8	homo sapien
5	120	100.0	2267	2	Q6SDP9	homo sapien
6	120	100.0	2296	2	Q6NOA6	homo sapien
7	120	100.0	2357	2	Q6BDT4	homo sapien
8	120	100.0	2386	1	FINC HUMAN	homo sapien
9	120	100.0	2444	2	Q6NO25	homo sapien
10	120	100.0	2477	2	Q6MZU5	homo sapien
11	119	99.2	2265	1	FINC BOVIN	bos taurus
12	115	95.8	141	2	Q90XQ2	ambystoma m
13	114	95.0	190	1	FINC NOTVI	noctophthalm
14	113	94.2	922	2	Q93405	brachydanio
15	113	94.2	2478	2	Q93406	brachydanio
16	112	93.3	2481	1	FINC XENLA	xenopus lae
17	112	93.3	2481	2	Q6QQA5	xenopus lae
18	110	91.7	2477	1	FINC MOUSE	mus musculus
19	107	89.2	2477	1	FINC RAT	rattus norv
20	96	80.0	2408	2	Q6JAN2	brachydanio
21	57.5	47.9	296	2	Q8CGJ7	mus musculus
22	53	44.2	768	2	Q8YE51	brucea m
23	53	44.2	859	2	Q8FPX25	brucea m
24	53	44.2	866	2	Q9FHI9	arabidopsis
25	52	43.3	322	2	Q20990	caenorhabdi
26	51.5	42.9	289	1	YMO1 CAEEL	caenorhabdi
27	51	42.5	143	2	Q9CZ77	mus musculus
28	50	41.7	140	2	Q99A64	bovine vira
29	50	41.7	140	2	Q99A65	bovine vira
30	50	41.7	140	2	Q99A69	bovine vira
31	50	41.7	140	2	Q99A70	bovine vira

```

GN Name=DKFZp686K139; Synonyms=DKFZp686F219;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansoerge W., Krieger S., Regiert T., Rittmüller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE46200.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON-TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;

Query Match 100.0%; Score 120; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 9,3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
DB 588 VNGRGWTCIAYSQLRDQCI 608

RESULT 3
Q6MZM7 PRELIMINARY; PRT; 2193 AA.
AC Q6MZM7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686O12165 (Fragment).
GN Name=DKFZp686O12165;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG The German Human cDNA Consortium;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640999; CAE46002.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
FT NON-TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;

Query Match 100.0%; Score 120; DB 2; Length 2193;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
DB 246 VNGRGWTCIAYSQLRDQCI 266

RESULT 4
Q68DP8 PRELIMINARY; PRT; 2240 AA.
AC Q68DP8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686H0342.
GN Name=DKFZp686H0342;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amvgdala;
RG The German cDNA Consortium;
RA Ottenwälder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
FT NON-TER 1
SQ SEQUENCE 2193 AA; 240641 MW; F876E93106540EF3 CRC64;

Query Match 100.0%; Score 120; DB 2; Length 2193;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
DB 246 VNGRGWTCIAYSQLRDQCI 266

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.59216 Seconds
(without alignments)
1369.066 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGSEWTCIAYSLRDQCI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	99.2	2265	1 FNBO	fibronectin - bovi
2	115	95.8	2386	1 FNBU	fibronectin precu
3	114	95.0	190	2 IS1279	fibronectin - east
4	112	93.3	2481	2 A43908	fibronectin - Afri
5	107	89.2	2477	2 S14428	fibronectin precu
6	53	44.2	768	2 A33505	alkaline phosphata
7	52	43.3	143	2 T28845	hypothetical prote
8	51.5	42.9	786	2 S44837	K02D10.1 protein -
9	49.5	41.2	420	2 S86614	hypothetical prote
10	48	40.0	718	2 A83282	probable TonB-depe
11	47.5	39.6	405	2 T12805	hypothetical prote
12	47	39.2	295	2 T49230	hypothetical prote
13	47	39.2	502	1 J02341	arginine decarboxy
14	47	39.2	612	2 T35430	probable long-chai
15	46	38.3	315	2 A64205	O-sialoglycoprotei
16	45.5	37.9	416	2 T10623	hypothetical prote
17	45	37.5	172	2 S16208	NADH2 dehydrogenas
18	45	37.5	355	2 B95213	ABC transporter, s
19	45	37.5	355	2 C98077	hypothetical prote
20	45	37.5	621	2 T15046	arginine decarboxy
21	44.5	37.1	366	2 S11363	actin modulator pr
22	44.5	37.1	780	2 T00366	hypothetical prote
23	44	36.7	147	2 S64252	probable membrane
24	44	36.7	256	2 C70687	hypothetical prote
25	44	36.7	319	2 S73421	o-sialoglycoprotei
26	44	36.7	437	1 FGHUG	fibrinogen gamma-A
27	44	36.7	453	1 FGHUGB	fibrinogen gamma-B
28	44	36.7	583	2 T25690	hypothetical prote
29	44	36.7	700	2 T27364	hypothetical prote

ALIGNMENTS

RESULT 1

FNBO

fibronectin - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A26452; B21165; A23292

R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A>Title: Complete primary structure of bovine plasma fibronectin.

A:Reference number: A26452; MUID:87054047; PMID:3780752

A:Accession: A26452

A:Molecule type: protein

A:Residues: 1-2265 <SKO>

A:Cross-references: UNIPROT:P07589

R:Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983

A>Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

A:Reference number: A21165; MUID:83221567; PMID:6304699

A:Accession: B21165

A:Molecule type: mRNA

A:Residues: 2170-2265 <KOR>

A:Cross-references: GB:K00800; NID:g163055; PIDN:AAA30521.2; PID:g5713333

R:Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottl

Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983

A>Title: Partial primary structure of bovine plasma fibronectin: three types of internal

A:Reference number: A23292; MUID:83117805; PMID:6218503

A:Accession: A23292

A:Molecule type: protein

A:Residues: 1-16, 'C', 18-20, 'S', 22-432, 447-463, 1367-1517, 1567-1673, 2062-2176, 'N', 2178-226

C:Comment: Cys-1201 and Cys-2015 have free sulphydryl groups.

C:Comment: The plasma fibronectin molecule consists of two chains, which are connected L

C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib

aling, and maintenance of cell shape.

C:Comment: Plasma fibronectin is synthesized by hepatocytes.

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe

C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu

F:21-241/Domain: fibrin and heparin binding <FBR>

F:21-56/Domain: fibronectin type I repeat homology <1F1>

F:66-104/Domain: fibronectin type I repeat homology <1F2>

F:110-148/Domain: fibronectin type I repeat homology <1F3>

F:155-194/Domain: fibronectin type I repeat homology <1F4>

F:200-239/Domain: fibronectin type I repeat homology <1F5>

F:277-577/Domain: collagen binding <CBR>

F:277-311/Domain: fibronectin type I repeat homology <1F6>

F:329-370/Domain: fibronectin type II repeat homology <2F1>

F:389-430/Domain: fibronectin type II repeat homology <2F2>

F:439-477/Domain: fibronectin type I repeat homology <1F7>

F:487-524/Domain: fibronectin type I repeat homology <1F8>

F:530-568/Domain: fibronectin type I repeat homology <1F9>

F:578-661/Domain: fibronectin type III repeat homology <FN3A>

F:688-770/Domain: fibronectin type III repeat homology <FN3B>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

genome polypeptide
hypothetical prote
probable membrane
maltose permease -
hypothetical prote
probable phosphogl
hypothetical prote
hypothetical prote
surface glycoprote
sugar permease hom
RNA helicase DbpA
surface glycoprote
genome polypeptide
probable dimethyl
protein F28C11.9
hypothetical prote

F:875-957/Domain: fibronectin type III repeat homology <FN3D>
 F:965-1046/Domain: fibronectin type III repeat homology <FN3E>
 F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
 F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
 F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>
 F:1326-1404/Domain: fibronectin type III repeat homology <FN3I>
 F:1410-1517/Domain: cell attachment <CAD>
 F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
 F:1493-1495/Region: cell attachment (R-G-D) motif
 F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
 F:1600-1870/Domain: heparin binding <HB2>
 F:1600-1882/Domain: fibronectin type III repeat homology <FN3L>
 F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>
 F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>
 F:1970-1972/Region: cell attachment (R-G-D) motif
 F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
 F:1985-2216/Domain: fibrin binding <FB2>
 F:2085-2124/Domain: fibronectin type I repeat homology <IF10>
 F:2130-2167/Domain: fibronectin type I repeat homology <IF11>
 F:2174-2209/Domain: fibronectin type I repeat homology <IF12>
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:21/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experimental
 F:21-47-45-56-64-92-104-110-138-136-148-155-184-182-194-200-229-227-239-277-304-302-3
 7-2155-2167-2174-2200-2198-2209/Disulfide bonds: #status predicted
 F:399-497-511-846-976-1213-1987/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F:1205-1692/Binding site: carbohydrate (Asn) (covalent) #status absent
 F:1943-1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F:2246/Disulfide bonds: interchain (to 2250) #status predicted
 F:2250/Disulfide bonds: interchain (to 2246) #status predicted
 F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 99.2%; Score 119; DB 1; Length 2265;

Best Local Similarity 95.2%; Pred. No. 5.7e-10;

Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21

DB 468 VNGRGWTCIAYSQLRDQCI 488

RESULT 2

FNHU fibronectin precursor [validated] - human

N:Alternate names: fibronectin splice form ED-A

C:Species: Homo sapiens (man)

C>Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C:Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R:Dean, D.C.; Bewluis, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A:Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A:Reference number: A26460; MUID:87175578; PMID:3031656

A:Accession: A26460

A:Molecule type: DNA

A:Residues: 1-49 <DEA>

A:Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA5337

R:Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A:Title: Evolution of the fibronectin gene.

A:Reference number: A26284; MUID:86111901; PMID:3003095

A:Accession: A26284

A:Molecule type: DNA

A:Residues: 1447-1540 <OLD>

A:Cross-references: GB:M12549; NID:g182688

A>Note: The authors translated the codon TTC for residue 1494 as Glu

R:Paolletta, G.; Henchcliffe, C.; Sebastiao, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A:Reference number: S00848; MUID:88233940; PMID:3375063

A:Accession: S03917

A:Molecule type: DNA

A:Residues: 1594-1767 'V', 1769-1783 <PAO>

A:Cross-references: EMBL:X07718; NID:g31402

A>Note: the authors translated the codon AAC for residue 1631 as Asp

R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 FEBS Lett. 207, 287-291, 1986
 A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:
 A:Reference number: A24854; MUID:87030929; PMID:3770201
 A:Accession: A24854
 A:Molecule type: DNA
 A:Residues: 1992-2147 <VIB>
 A:Cross-references: GB:X04530; NID:g31436
 R:Gutman, A.; Yamada, K.M.; Kornblitt, A.
 FEBS Lett. 207, 145-148, 1986
 A:Title: Human fibronectin is synthesized as a pre-propolypeptide.
 A:Reference number: A24476; MUID:87030890; PMID:3770189
 A:Accession: A24476
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-14, 'Q', 16-38 <GUT>
 R:Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A:Title: Primary structure of human fibronectin: differential splicing may generate at l
 A:Reference number: A91008; MUID:85284965; PMID:2992939
 A:Accession: A91008
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
 A:Cross-references: GB:X02761
 R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypepti
 A:Reference number: A93529; MUID:84272258; PMID:6462919
 A:Accession: A93529
 A:Molecule type: mRNA
 A:Residues: 973-2080; 2112-2386 <KO2>
 A:Cross-references: GB:X00739
 R:Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
 A:Reference number: A21011; MUID:83290929; PMID:6688418
 A:Accession: A21011
 A:Molecule type: mRNA
 A:Residues: 1434-1537 <OL2>
 A:Cross-references: GB:K00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
 R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2898-2704, 1985
 A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
 A:Reference number: A90495; MUID:85280409; PMID:2992573
 A:Accession: A90495
 A:Molecule type: mRNA
 A:Residues: 1594-2386 <BER>
 A:Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
 R:Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
 FEBS Lett. 186, 31-34, 1985
 A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
 A:Reference number: A22245; MUID:85231203; PMID:2989004
 A:Accession: A22245
 A:Molecule type: mRNA
 A:Residues: 1948-2067 <UME>
 A:Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
 A:Accession: B22245
 A:Molecule type: mRNA
 A:Residues: 1975-1991, 2017-2039 <UM2>
 A:Cross-references: GB:M27590
 R:Seikiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
 Biochemistry 25, 4936-4941, 1986
 A:Title: Human liver fibronectin complementary DNAs: identification of two different mē
 A:Reference number: I52394; MUID:87026578; PMID:3021206
 A:Accession: I65273
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
 A:Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
 R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 7.1098 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGWETCIAYSQLRDQCI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	343	9 US-09-934-706-1	Sequence 1, Appli
2	120	100.0	400	9 US-09-934-706-5	Sequence 5, Appli
3	120	100.0	501	9 US-09-934-706-4	Sequence 4, Appli
4	120	100.0	642	16 US-10-741-601-354	Sequence 354, App
5	120	100.0	642	17 US-10-741-600-1066	Sequence 1066, Ap
6	120	100.0	657	16 US-10-741-601-359	Sequence 359, App
7	120	100.0	657	17 US-10-741-600-1072	Sequence 1072, Ap
8	120	100.0	984	16 US-10-741-601-356	Sequence 356, App
9	120	100.0	984	17 US-10-741-600-1069	Sequence 1069, Ap
10	120	100.0	1173	18 US-10-450-763-52634	Sequence 52634, A
11	120	100.0	2220	15 US-10-236-392-4	Sequence 4, Appli

12	120	100.0	2296	16 US-10-741-601-363	Sequence 363, App
13	120	100.0	2296	17 US-10-741-600-1075	Sequence 1075, Ap
14	120	100.0	2320	14 US-10-279-733-8	Sequence 8, Appli
15	120	100.0	2320	15 US-10-236-392-2	Sequence 2, Appli
16	120	100.0	2328	14 US-10-171-311-64	Sequence 64, Appli
17	120	100.0	2328	15 US-10-236-031B-70	Sequence 70, Appli
18	120	100.0	2328	15 US-10-374-979-98	Sequence 98, Appli
19	120	100.0	2328	15 US-10-182-936A-98	Sequence 98, Appli
20	120	100.0	2328	16 US-10-477-238A-677	Sequence 677, App
21	120	100.0	2328	16 US-10-680-287A-677	Sequence 677, App
22	120	100.0	2328	17 US-10-477-173-677	Sequence 104, App
23	120	100.0	2355	15 US-10-144-194A-104	Sequence 235, App
24	120	100.0	2355	15 US-10-360-101-235	Sequence 3, Appli
25	120	100.0	2355	15 US-10-447-161-3	Sequence 94, Appli
26	120	100.0	2355	16 US-10-734-564-94	Sequence 357, App
27	120	100.0	2355	16 US-10-741-601-357	Sequence 366, App
28	120	100.0	2355	16 US-10-741-601-366	Sequence 104, App
29	120	100.0	2355	16 US-10-491-566-104	Sequence 1067, Ap
30	120	100.0	2355	17 US-10-741-600-1067	Sequence 1078, Ap
31	120	100.0	2355	17 US-10-741-600-1078	Sequence 147, App
32	120	100.0	2355	17 US-10-852-335A-147	Sequence 436, App
33	120	100.0	2355	18 US-10-287-436A-436	Sequence 1137, Ap
34	120	100.0	2355	18 US-10-287-436A-1137	Sequence 28, Appli
35	120	100.0	2386	20 US-11-040-130-28	Sequence 1, Appli
36	120	100.0	2386	10 US-09-961-403-1	Sequence 360, App
37	120	100.0	2386	16 US-10-741-601-360	Sequence 206, App
38	120	100.0	2386	16 US-10-788-792-206	Sequence 32, Appli
39	120	100.0	2386	16 US-10-618-281-32	Sequence 1071, Ap
40	120	100.0	2386	17 US-10-741-600-1071	Sequence 59, Appli
41	120	100.0	2386	17 US-10-868-577A-59	Sequence 4, Appli
42	120	100.0	2386	18 US-10-485-758-4	Sequence 9, Appli
43	120	100.0	2386	18 US-10-485-758-9	Sequence 52638, A
44	120	100.0	2474	18 US-10-450-763-52638	Sequence 4, Appli
45	96	80.0	228	15 US-10-344-634-4	

ALIGNMENTS

RESULT 1
US-09-934-706-1
; Sequence 1, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified Human
; OTHER INFORMATION: Fibronectin Collagen-Binding Domain
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: CONFLICT
; LOCATION: (69)
; NAME/KEY: CONFLICT
; LOCATION: (125)
US-09-934-706-1

Query Match 100.0%; Score 120; DB 9; Length 343;

Best Local Similarity 100.0%; Pred. No. 3.9e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 2
US-09-934-706-5
; Sequence 5, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Hybrid
; OTHER INFORMATION: Polypeptide of Human Fibronectin Collagen-Binding
; OTHER INFORMATION: Domain and Human Epidermal Growth Factor
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: PEPTIDE
; LOCATION: (343)..(347)
; OTHER INFORMATION: /note=" enterokinase recognition sequence"
; NAME/KEY: PEPTIDE
; LOCATION: (348)..(400)
; OTHER INFORMATION: /note="human epidermal growth factor"
US-09-934-706-5

Query Match 100.0%; Score 120; DB 9; Length 400;
Best Local Similarity 100.0%; Pred. No. 4.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 3
US-09-934-706-4
; Sequence 4, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Hybrid
; OTHER INFORMATION: Polypeptide of Human Fibronectin Collagen-Binding
; OTHER INFORMATION: Domain and Human Basic Fibroblast Growth Factor
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN

```

```

; LOCATION: (2)..(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: PEPTIDE
; LOCATION: (343)..(347)
; OTHER INFORMATION: /note="enterokinase recognition sequence"
; NAME/KEY: PEPTIDE
; LOCATION: (348)..(501)
; OTHER INFORMATION: /note="human fibroblast growth factor"
US-09-934-706-4

Query Match 100.0%; Score 120; DB 9; Length 501;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 4
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 100.0%; Score 120; DB 16; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 484 VNGRGWTCIAYSQLRDQCI 504

RESULT 5
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match 100.0%; Score 120; DB 17; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
   |||||
Db 484 VNGRGWTCIAYSQLRDQCI 504

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.22353 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120

Sequence: 1 VGNRGWTCIAYSQLRDQCI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	2324	1 US-08-283-857-1	Sequence 1, Appli
2	120	100.0	2324	5 PCT-US95-09819-1	Sequence 1, Appli
3	120	100.0	2327	6 5455158-1	Patent No. 5455158
4	120	100.0	2327	6 5455158-1	Patent No. 5455158
5	120	100.0	2386	4 US-09-961-403-1	Sequence 1, Appli
6	120	100.0	2446	2 US-08-551-356-2	Sequence 2, Appli
7	120	100.0	2446	5 PCT-US93-12887-2	Sequence 2, Appli
8	115	95.8	2231	1 US-08-153-799-16	Sequence 16, Appli
9	115	95.8	2386	2 US-09-016-366A-12	Sequence 20, Appli
10	102	85.0	48	2 US-08-982-597A-20	Sequence 12, Appli
11	102	85.0	48	3 US-09-136-218-20	Sequence 20, Appli
12	50	41.7	115	4 US-09-902-540-12185	Sequence 12185, A
13	48	40.0	723	4 US-09-252-991A-18279	Sequence 18279, A
14	46	38.3	231	4 US-09-328-352-6328	Sequence 6328, Ap
15	45	37.5	355	4 US-09-583-110-3026	Sequence 3026, Ap
16	45	37.5	729	4 US-09-543-681A-8257	Sequence 8257, Ap
17	44	36.7	20	4 US-09-664-945-4	Sequence 4, Appli
18	44	36.7	27	4 US-09-664-945-65	Sequence 65, Appli
19	44	36.7	122	4 US-09-513-999C-7822	Sequence 7822, Ap
20	44	36.7	147	4 US-09-538-092-334	Sequence 324, App
21	44	36.7	206	4 US-09-902-540-16467	Sequence 16467, A
22	44	36.7	225	4 US-09-270-767-60597	Sequence 60597, A
23	44	36.7	302	4 US-09-270-767-45103	Sequence 45103, A
24	44	36.7	449	4 US-09-949-016-9614	Sequence 9614, Ap
25	44	36.7	449	4 US-09-949-016-9615	Sequence 9615, Ap
26	44	36.7	453	1 US-08-206-176-6	Sequence 6, Appli
27	44	36.7	454	3 US-08-434-099A-27	Sequence 27, Appli

28 44 36.7 454 3 US-08-434-099A-28 Sequence 28, Appli
29 44 36.7 465 4 US-09-949-016-8583 Sequence 8583, Ap
30 44 36.7 465 4 US-09-949-016-8584 Sequence 8584, Ap
31 43.5 36.2 627 4 US-09-949-016-9447 Sequence 9447, Ap
32 43.5 36.2 761 4 US-09-949-016-6880 Sequence 6880, Ap
33 43.5 36.2 1068 4 US-09-248-796A-16119 Sequence 16119, A
34 43 35.8 10 3 US-08-456-466-27 Sequence 27, Appli
35 43 35.8 54 4 US-09-621-976-4577 Sequence 4577, Ap
36 43 35.8 165 4 US-09-270-767-37136 Sequence 37136, A
37 43 35.8 165 4 US-09-270-767-52353 Sequence 52353, A
38 43 35.8 187 4 US-09-640-211A-821 Sequence 821, App
39 43 35.8 251 4 US-09-540-236-3406 Sequence 3406, Ap
40 43 35.8 423 1 US-08-445-746-2 Sequence 2, Appli
41 43 35.8 423 3 US-09-008-722-2 Sequence 2, Appli
42 43 35.8 431 4 US-09-248-796A-18354 Sequence 18354, A
43 42.5 35.4 564 4 US-09-328-352-7905 Sequence 7905, Ap
44 42 35.0 43 2 US-08-982-597A-21 Sequence 21, Appli
45 42 35.0 43 3 US-09-136-218-21 Sequence 21, Appli

ALIGNMENTS

RESULT 1
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5792742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283.857
; FILING DATE: 01-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,134
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-283-857-1

Query Match 100.0%; Score 120; DB 1; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2e-09;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
Db 468 VNGRGWTCIAYSQLRDQCI 488

RESULT 2

PCT-US95-09819-1
; Sequence 1, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-09819-1

Query Match 100.0%; Score 120; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
Db 468 VNGRGWTCIAYSQLRDQCI 488

RESULT 3

5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990

; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
; SEQ ID NO:1:
; LENGTH: 2327
5455158-1

Query Match 100.0%; Score 120; DB 6; Length 2327;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
Db 471 VNGRGWTCIAYSQLRDQCI 491

RESULT 4

5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
; SEQ ID NO:1:
; LENGTH: 2327
5455158-1

Query Match 100.0%; Score 120; DB 6; Length 2327;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VNGRGWTCIAYSQLRDQCI 21
Db 471 VNGRGWTCIAYSQLRDQCI 491

RESULT 5

US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

Query Match 100.0%; Score 120; DB 4; Length 2386;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.76863 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGWETCIAYSQLRDQCI 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: geneseqp1980s.*
2: geneseqp1980s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	100.0	21	2 AAY28906	AAY28906 MSF 1-1alp
2	120	100.0	343	3 AAB08505	AAB08505 Amino aci
3	120	100.0	343	5 ABB07961	ABB07961 Modified
4	120	100.0	400	3 AAB08509	AAB08509 Hybrid of
5	120	100.0	461	5 ABB07964	ABB07964 Human fib
6	120	100.0	473	5 ABB07965	ABB07965 Human fib
7	120	100.0	501	3 AAB08508	AAB08508 Hybrid of
8	120	100.0	642	8 ADQ39403	ADQ39403 Human myo
9	120	100.0	642	8 ADR67316	ADR67316 Human bla
10	120	100.0	642	8 ADS17489	ADS17489 Amino aci
11	120	100.0	642	8 ADR97658	ADR97658 Human fib
12	120	100.0	657	8 ADQ39409	ADQ39409 Human myo
13	120	100.0	660	2 AAY28901	AAY28901 Human mig
14	120	100.0	984	8 ADQ39406	ADQ39406 Human myo
15	120	100.0	1173	4 ABG22275	ABG22275 Novel hum
16	120	100.0	1179	8 ADP75952	ADP75952 Human min
17	120	100.0	1223	8 ADP30365	ADP30365 Human sec
18	120	100.0	1359	8 ADP75957	ADP75957 Human leu
19	120	100.0	2182	8 ADR66462	ADR66462 Human pro
20	120	100.0	2182	8 ADR66120	ADR66120 Human pro
21	120	100.0	2220	6 ABO01289	ABO01289 Human pro
22	120	100.0	2220	8 ADN95950	ADN95950 Human NOV
23	120	100.0	2265	4 AAM38647	AAM38647 Human pol
24	120	100.0	2266	6 ABR40124	ABR40124 Human cel
25	120	100.0	2296	8 ADQ39412	ADQ39412 Human myo

ALIGNMENTS

RESULT 1

AAY28906
ID AAY28906 standard; peptide; 21 AA.

XX AC AAY28906;

DT 21-SEP-1999 (first entry)

DE MSF 1-alpha peptide epitope.

XX Migration stimulatory factor; MSF; cell-migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.

OS Synthetic.

OS Homo sapiens.

XX WO9931233-Al.

XX 24-JUN-1999.

PF 15-DEC-1998; 98WO-GB003766.

PR 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI, 1999-430039/36.

XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX Claim 18; Page 57; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28902-906 represent peptide epitopes of MSF against which monoclonal antibodies that are specific to MSF and do not cross-react with fibronectin are raised

XX Sequence 21 AA;

Query Match 100.0%; Score 120; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;

26	120	100.0	2320	6 AAE37107	AAE37107 Human fib
27	120	100.0	2320	6 ABR42588	ABR42588 Human fib
28	120	100.0	2320	6 ABO01288	ABO01288 Human pro
29	120	100.0	2320	8 ADR90519	ADR90519 Human ful
30	120	100.0	2324	5 AAU74674	AAU74674 Human fib
31	120	100.0	2324	5 AAE23651	AAE23651 Human fib
32	120	100.0	2327	1 AAP70373	AAP70373 Human pro
33	120	100.0	2327	2 AAR15468	AAR15468 Human fib
34	120	100.0	2328	4 AAG68182	AAG68182 Fibronect
35	120	100.0	2328	6 ABU07486	ABU07486 Protein d
36	120	100.0	2328	6 ABR41106	ABR41106 Human fib
37	120	100.0	2328	6 ABR92078	ABR92078 Human cer
38	120	100.0	2328	7 ADB70378	ADB70378 Fibronect
39	120	100.0	2328	7 ADB98726	ADB98726 Human fib
40	120	100.0	2328	7 ADE82522	ADE82522 Human pro
41	120	100.0	2328	8 ADJ37157	ADJ37157 Human mal
42	120	100.0	2330	4 AAM38646	AAM38646 Human pol
43	120	100.0	2355	4 AAM38649	AAM38649 Human pol
44	120	100.0	2355	6 ABR58335	ABR58335 NM_00202
45	120	100.0	2355	6 ABR58335	ABR58335 NM_00202

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNGRGWTCIAYSQLRDQCI 21
DB 1 VNGRGWTCIAYSQLRDQCI 21

RESULT 2
AAB08505
ID AAB08505 standard; protein; 343 AA.
XX AAB08505;
XX 20-DEC-2000 (first entry)
XX Amino acid sequence of modified fibronectin collagen-binding domain.
XX Fibronectin; collagen-binding domain; sustained release; gene therapy;
KW physiologically active polypeptide; topical retention;
KW tissue regeneration.
XX Synthetic.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Misc-difference 1 /note= "Met added for efficient expression"
FT WO200049159-A1.
PN 24-AUG-2000.
XX 21-FEB-2000; 2000WO-JP000964.
XX 19-FEB-1999; 99JP-00041913.
PR 01-NOV-1999; 99JP-00311364.
XX (TERU) TERUMO CORP.
XX Ishikawa T, Kitajima T;
XX WPI; 2000-565375/52.
DR N-PSDB; AAA64263.
XX Collagen-binding active polypeptide for use in an agent for enabling
PT topical retention or sustained release of a physiologically active
PT peptide or physiological activity-imparting agent comprises a fibronectin
PT peptide.
XX Disclosure; Page 113-115; 135pp; English.
XX The present sequence represents a modified fibronectin collagen-
XX binding domain. The protein is used to construct a collagen-binding
CC physiologically active polypeptide. This polypeptide comprises a peptide
CC from fibronectin ligated to a physiologically active peptide. The
CC polypeptides are used in an agent for enabling topical retention or
CC sustained release of a physiologically active peptide or physiological
CC activity-imparting agent. They may be used in gene therapy and in tissue
XX regeneration
XX Sequence 343 AA;

Query Match 100.0%; Score 120; DB 3; Length 343;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNGRGWTCIAYSQLRDQCI 21
DB 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 3
AAB07961

ID AAB07961 standard; protein; 343 AA.
XX AAB07961;
XX 12-AUG-2002 (first entry)
XX Modified human fibronectin collagen-binding domain.
XX Osteogenesis; drug delivery system; DDS; collagen; osteopathic; human;
KW fibronectin.
XX Homo sapiens.
OS Synthetic.
XX JP2002058485-A.
XX 26-FEB-2002.
XX 16-AUG-2000; 2000JP-00246744.
XX 16-AUG-2000; 2000JP-00246744.
XX (TERU) TERUMO CORP.
XX WPI; 2002-438612/47.
DR N-PSDB; ABL41020.
XX Novel osteogenesis stimulating fused protein having collagen avidity used
PT as an osteogenesis stimulator and a localizing agent.
PT Disclosure; Page 17-18; 30pp; Japanese.
XX The invention provides an osteogenesis stimulating fused protein (I) for
CC a drug delivery system (DDS) of osteogenetic factor, having collagen
CC avidity and polypeptides homologous to collagen avidity domain or its
CC modified peptides. (I) is used for stimulation of osteogenesis, a
CC localizing agent and a slow releasing agent for a drug delivery system.
CC The present sequence represents a modified human fibronectin collagen-
CC binding domain
XX Sequence 343 AA;

Query Match 100.0%; Score 120; DB 5; Length 343;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VNGRGWTCIAYSQLRDQCI 21
DB 210 VNGRGWTCIAYSQLRDQCI 230

RESULT 4
AAB08509
ID AAB08509 standard; protein; 400 AA.
XX AAB08509;
XX 12-SEP-2003 (revised)
DT 20-DEC-2000 (first entry)
XX Hybrid of fibronectin collagen-binding domain/epidermal growth factor.
XX Fibronectin; collagen-binding domain; sustained release; gene therapy;
KW physiologically active polypeptide; topical retention; chimera;
KW tissue regeneration; epidermal growth factor.
XX Homo sapiens.
OS Homo sapiens.
OS Chimeric.
XX Key Location/Qualifiers
FH Protein 2..341
FT /note= "human fibronectin collagen-binding domain"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 436.745 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGGEWTCIAYSLRDQCI 21

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US9581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1 1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	207	1	AL710586 DKFZp686N
c 2	120	100.0	242	5	BQ340610 PMO-NN025
3	120	100.0	304	1	AA852107 NHTBCae10
4	120	100.0	321	4	BI036180 CM3-NT026
c 5	120	100.0	369	5	BQ340603 PMO-NN025
6	120	100.0	373	2	BE817152 PM3-BN025
c 7	120	100.0	380	5	BQ340568 PMO-NN025
c 8	120	100.0	385	5	BQ340613 PMO-NN025
9	120	100.0	386	7	CN419560 170005321

10	120	100.0	407	7	CN419462	170005322
11	120	100.0	436	2	BF868748	CMO-ET009
12	120	100.0	461	7	CF128990	UI-HF-BS0
13	120	100.0	480	7	CN419466	170005319
c 14	120	100.0	497	4	EG000888	ILO-GN024
15	120	100.0	503	7	CN419452	170004241
c 16	120	100.0	525	2	AW853981	RC3-CT025
17	120	100.0	530	1	AU141012	AU141012
18	120	100.0	534	5	BK642843	DKFZp781M
c 19	120	100.0	536	4	EG000869	ILO-GN024
c 20	120	100.0	542	5	BQ340573	PMO-NN025
21	120	100.0	592	7	CN419590	170004247
22	120	100.0	613	6	CD613819	56080991J
23	120	100.0	618	7	CN481692	CN41692 hwi1a08.Y
24	120	100.0	618	7	CN482377	hwi19f03.Y
25	120	100.0	632	7	CN482351	hwi19c10.Y
26	120	100.0	635	5	BQ312550	PMO-BN017
27	120	100.0	645	7	CN419499	170004551
28	120	100.0	653	7	CN419474	170004251
29	120	100.0	655	7	CN419546	170005328
30	120	100.0	657	7	CN419503	170004706
31	120	100.0	667	1	AU140910	AU140910
c 32	120	100.0	670	5	BQ620723	UI-H-FL1-
33	120	100.0	673	1	AU140450	AU140450
34	120	100.0	684	1	AU134619	AU134619
35	120	100.0	695	7	CN419477	170004247
36	120	100.0	699	1	AU140889	AU140889
37	120	100.0	699	7	CN419497	170004240
38	120	100.0	702	7	CN419472	170005326
39	120	100.0	705	1	AU140899	AU140899
40	120	100.0	707	1	AU140730	AU140730
41	120	100.0	723	7	CN419622	170005319
42	120	100.0	730	1	AU140834	AU140834
43	120	100.0	731	1	AU140526	AU140526
44	120	100.0	732	1	AU140747	AU140747
45	120	100.0	734	6	CD613786	56002834H

ALIGNMENTS

RESULT 1
AL710586
LOCUS DKFZp686N066.1 686 (synonym: hicc3) Homo sapiens linear EST 04-SEP-2003
DEFINITION DKFZp686N066.5', mRNA sequence.
ACCESSION AL710586
VERSION AL710586
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 207)
AUTHORS Ansong, W., Wirkner, U., Mewes, W., Weill, B. and Wiemann, S.
TITLE EST (Ansong, W., Wirkner, U., Mewes, W., Weill, B. and Wiemann, S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686N066) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..207
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="taxon:9606"
/clone="DKFZp686N066"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hicc3)"
/notes="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3 98e-08 Length: 207
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

```

US-09-581-651D-9 (1-21) x AL710586 (1-207)

```

Qy 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
|||||
Db 54 GTTGGGAATGCTGCTGGGAATGGATGATGCTACTGCGAGCTTCGAGATCAGTGC 113
|||||
Qy 21 Ile 21
|||
Db 114 ATT 116

```

RESULT 2

```

BQ340610/C 242 bp mRNA linear EST 20-MAY-2002
LOCUS BQ340610
DEFINITION PMO-NN0258-220501-001-e02 NN0258 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ340610
VERSION BQ340610.1 GI:21001672
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

```

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

CONTACT Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PMO&t2=PMO-NN0258-

220501-001-e02&t3=2001-05-22&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 16

High quality sequence stop: 241.

FEATURES

Location/Qualifiers

1..242

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="NN0258"

/note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

```

Alignment Scores:
Pred. No.: 4.75e-08 Length: 242
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

```

US-09-581-651D-9 (1-21) x BQ340610 (1-242)

```

Qy 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
|||||
Db 155 GTTGGGAATGCTGCTGGGAATGGATGATGCTACTGCGAGCTTCGAGATCAGTGC 96
|||||
Qy 21 Ile 21
|||
Db 95 ATT 93

```

RESULT 3

```

AA852107 304 bp mRNA linear EST 20-JUN-2002
LOCUS AA852107
DEFINITION NHTBCae10c08r1 Normal Human Trabecular Bone Cells Homo sapiens cDNA
clone NHTBCae10c08, mRNA sequence.
ACCESSION AA852107
VERSION AA852107.1 GI:2940700
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
SGAP: The Skeletal Genome Anatomy Project
Unpublished (1997)
Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1367, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
Seq primer: M13 Reverse.

```

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..304

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NHTBCae10c08"

/sex="Female"

/tissue_type="Bone"

/cell_type="Trabecular Bone Cells"

/lab_host="SURE"

/clone_lib="Normal Human Trabecular Bone Cells"

/note="Organ: Hip; Vector: pBluescript; Site 1: EcoRI;

Library constructed by Dr. Marian Young and Dr. Pamela

Gehron Robey (NIDCR)"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservativeness:

Mismatches:

Indels:

Gaps:

Mismatches:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 107.827 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGWTCIAYSQLRDQCI 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 413490567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-LOOPEXT=0 -GOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09581651@cgn1.1.1041@runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*
1: /cgn2_1/pubna/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	100.0	358	18	US-10-242-535A-30516
2	120	100.0	358	19	US-10-085-783A-30516
3	120	100.0	1053	9	US-09-934-706-8
4	120	100.0	1189	15	US-10-198-846-11556
5	120	100.0	1224	9	US-09-934-706-16
6	120	100.0	1527	9	US-09-934-706-15
7	120	100.0	2127	17	US-10-210-120-49
8	120	100.0	2127	22	US-10-956-157-4288
9	120	100.0	2127	24	US-10-309-035-49
10	120	100.0	2443	20	US-10-741-601-70
11	120	100.0	2443	20	US-10-741-600-238
12	120	100.0	2488	20	US-10-741-601-75
13	120	100.0	2488	22	US-10-741-600-244
14	120	100.0	3522	24	US-10-450-763-22266
15	120	100.0	4295	18	US-10-144-194A-51
16	120	100.0	4295	21	US-10-491-566-51
17	120	100.0	6510	20	US-10-741-601-72
18	120	100.0	6510	22	US-10-741-600-241
19	120	100.0	6988	19	US-10-236-392-1
20	120	100.0	7361	19	US-10-236-392-3
21	120	100.0	7677	22	US-10-956-157-4995
22	120	100.0	7679	22	US-10-831-704-38
23	120	100.0	7680	9	US-09-964-824A-574
24	120	100.0	7680	15	US-10-171-311-63
25	120	100.0	7680	18	US-10-236-031B-69
26	120	100.0	7680	18	US-10-374-979-75
27	120	100.0	7680	19	US-10-182-936A-75
28	120	100.0	7680	19	US-10-641-643-1289
29	120	100.0	7680	20	US-10-717-597-222
30	120	100.0	7680	20	US-10-788-792-79
31	120	100.0	7680	21	US-10-477-238A-654
32	120	100.0	7680	21	US-10-580-287A-654
33	120	100.0	7680	22	US-10-278-698-88
34	120	100.0	7680	22	US-10-278-698-603
35	120	100.0	7680	22	US-10-843-641A-5877
36	120	100.0	7680	22	US-10-477-173-654
37	120	100.0	7680	22	US-10-852-335A-52
38	120	100.0	7705	18	US-10-447-161-4
39	120	100.0	7795	16	US-10-084-817-2
40	120	100.0	7823	20	US-10-741-601-77
41	120	100.0	7823	22	US-10-741-600-245
42	120	100.0	7848	20	US-10-741-601-78
43	120	100.0	7848	22	US-10-741-600-246
44	120	100.0	7867	14	US-10-098-841-6
45	120	100.0	7935	20	US-10-741-601-74

ALIGNMENTS

RESULT 1
US-10-242-535A-30516
; Sequence 30516, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242.535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30516
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-30516

Alignment Scores:
Pred. No.: 1.49e-10 Length: 358
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-581-651D-9 (1-21) x US-10-242-535A-30516 (1-358)

QY 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 199 GTTGGGAATGGTCGTGGGAATGGACATGTCATTGCTACTCGCAGCTTCGAGATCAGTGC 258
QY 21 Ile 21
Db 259 ATT 261

RESULT 2

US-10-085-783A-30516
; Sequence 30516, Application US/10085783A
; Publication No. US20040037841A1

; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30516
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-30516

Alignment Scores:
Pred. No.: 1.49e-10 Length: 358
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-09-581-651D-9 (1-21) x US-10-085-783A-30516 (1-358)

QY 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 199 GTTGGGAATGGTCGTGGGAATGGACATGTCATTGCTACTCGCAGCTTCGAGATCAGTGC 258
QY 21 Ile 21
Db 259 ATT 261

RESULT 3

US-09-934-706-8
; Sequence 8, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Tetumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 8
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified Human
; OTHER INFORMATION: Fibronectin Collagen-Binding Domain
; NAME/KEY: conflict
; LOCATION: (109)
; NAME/KEY: conflict
; LOCATION: (206)
; NAME/KEY: conflict
; LOCATION: (270)
; NAME/KEY: conflict
; LOCATION: (374)
; NAME/KEY: conflict
; LOCATION: (681)
US-09-934-706-8

Alignment Scores:

Pred. No.: 5.06e-10 Length: 1053
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-9 (1-21) x US-09-934-706-8 (1-1053)

QY 1 ValGlyAenGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 643 GTTGGGAATGGTCGTGGGAATGGACATGTCATTGCTACTCGCAGCTTCGAGATCAGTGC 702
QY 21 Ile 21
Db 703 ATT 705

RESULT 4

US-10-198-846-11556/c
; Sequence 11556, Application US/10198846
; Publication No. US20030099974A1

; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11556
; LENGTH: 1189
; TYPE: DNA

Qy 21 Ile 21
Db 1475 ATT 1477

RESULT 2
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: g31396
US-09-023-655-1289

Alignment Scores:
Pred. No.: 6.96e-09 Length: 7680
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-9 (1-21) x US-09-023-655-1289 (1-7680)

Qy 1 ValGlyAanGlyArgGlyGluTrpThrCysIleAlaTySerGlnLeuArgAspGlnCys 20
Db 1415 GTTGGGAATGTCGTGGGGAATGGACATGCTACTGCGCAACTTCGAGATCAGTGC 1474

Qy 21 Ile 21
Db 1475 ATT 1477

RESULT 3

PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US95-09819-6

Alignment Scores:
Pred. No.: 6.96e-09 Length: 7680
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-581-651D-9 (1-21) x PCT-US95-09819-6 (1-7680)

Qy 1 ValGlyAanGlyArgGlyGluTrpThrCysIleAlaTySerGlnLeuArgAspGlnCys 20
Db 1415 GTTGGGAATGTCGTGGGGAATGGACATGCTACTGCGCAACTTCGAGATCAGTGC 1474

Qy 21 Ile 21
Db 1475 ATT 1477

RESULT 4
US-08-259-569-16
; Sequence 16, Application US/08259569
; Patent No. 5679320
; GENERAL INFORMATION:
; APPLICANT: Vogel, Tikva
; APPLICANT: Levanon, Avigdor
; APPLICANT: Werber, Moshe
; APPLICANT: Guy, Rachel
; APPLICANT: Panet, Amos
; APPLICANT: Hartman, Jacob
; APPLICANT: Shaked, Hadassa
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 68.8471 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-9

Perfect score: 120

Sequence: 1 VGNRGSEWTCIAYSLRDQCI 21

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092222_28772/app_query.fasta.1.1834
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=text -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651_QCGN 1 1052 @runat_07112005_092222_28772 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	120	100.0	1051	13 ADS19042	Ads19042 Human fib
2	120	100.0	1053	3 AA64263	Aa64263 cDNA enco
3	120	100.0	1053	6 ABL41020	Ab141020 Modified
4	120	100.0	1189	11 ACN90406	Acn90406 Breast ca
5	120	100.0	1224	3 AA64271	Aa64271 Fibronect

6	120	100.0	1407	6 ABL41027	Ab141027 Human fib
7	120	100.0	1443	6 ABL41028	Ab141028 Human fib
8	120	100.0	1527	3 AA64270	Aa64270 Fibronect
9	120	100.0	1929	13 ADS17488	Adsl7488 Nucleotid
10	120	100.0	1929	13 AD97657	Ad97657 Human fib
11	120	100.0	2127	10 ADD18477	Add18477 Human pro
12	120	100.0	2147	2 AAX81299	Aax81299 Human mlg
13	120	100.0	2358	13 AD67201	Ad67201 Human bla
14	120	100.0	2443	13 AD38575	Adq38575 Human SNP
15	120	100.0	2488	13 AD38581	Adq38581 Human SNP
16	120	100.0	3522	5 AAS86462	Aas86462 DNA enco
17	120	100.0	3540	12 ADP75954	Adp75954 Human min
18	120	100.0	3669	12 ADP29134	Adp29134 Human sec
19	120	100.0	4080	12 ADP75958	Adp75958 Human leu
20	120	100.0	4295	8 ACC72037	Acc72037 BC00770 g
21	120	100.0	4860	3 AAA35009	Aaa35009 Human ade
22	120	100.0	6510	13 AD38578	Adq38578 Human SNP
23	120	100.0	6815	10 ABX74443	Abx74443 Human cDN
24	120	100.0	6988	9 ACD06169	Acd06169 Human cDN
25	120	100.0	7020	13 ADR90518	Adr90518 Human cDN
26	120	100.0	7242	13 ADR66637	Adr66637 Human pro
27	120	100.0	7242	13 ADR65953	Adr65953 Human pro
28	120	100.0	7361	9 ACD06170	Acd06170 Human cDN
29	120	100.0	7361	12 ADN95949	Adn95949 Human NOV
30	120	100.0	7550	8 ACC00412	Acc00412 Human cel
31	120	100.0	7677	13 ACN37820	Acn37820 Tumour-as
32	120	100.0	7679	10 ADB31322	Adb31322 Testoster
33	120	100.0	7680	2 AAT17551	Aat17551 Human fib
34	120	100.0	7680	3 AAF21131	Aaf21131 Human low
35	120	100.0	7680	5 ABA82689	Aba82689 Fibronect
36	120	100.0	7680	6 ABL67540	Ab167540 Thyroid c
37	120	100.0	7680	6 ABT11082	Abt11082 Human bre
38	120	100.0	7680	8 ACF03878	Acf03878 Human fib
39	120	100.0	7680	8 ABX10391	Abx10391 DNA enco
40	120	100.0	7680	8 ACC46009	Acc46009 Human ear
41	120	100.0	7680	8 ACF12859	Acf12859 Human ear
42	120	100.0	7680	8 ACA64817	Ac64817 Human fib
43	120	100.0	7680	8 ACA64819	Ac64819 Human fib
44	120	100.0	7680	9 ADB70377	Adb70377 Fibronect
45	120	100.0	7680	10 ADB98703	Adb98703 Human fib

ALIGNMENTS

RESULT 1
ADS19042
ID ADS19042 standard; DNA; 1051 BP.

XX AC ADS19042;

XX DT 16-DEC-2004 (first entry)

XX DE Human fibronectin collagen binding domain DNA SeqID 3.

XX DE human; ds; hepatocyte growth factor; HGF; collagen binding domain;
XX KW fibronectin; artificial blood vessel; stent; tissue regeneration;
XX KW vulnary.

XX OS Homo sapiens.

XX PN JP2004269423-A.

XX PD 30-SEP-2004.

XX PF 07-MAR-2003; 2003JP-00062169.

XX PR 07-MAR-2003; 2003JP-00062169.

XX PA (TERU) TERUMO CORP.

XX DR WPI; 2004-682709/67.

XX PT Hepatocyte growth factor HGF derivative for inducing angiogenesis,

PT comprises fusion protein containing HGF polypeptide and polypeptide other
PT than HGF, and polypeptide having collagen binding property connected
PT directly or through intron.

XX Example 1; SEQ ID NO 3; 51pp; Japanese.

XX This invention relates to a novel hepatocyte growth factor (HGF)
CC derivative. Specifically, it refers to a fusion protein containing an HGF
CC polypeptide and the collagen binding domain of fibronectin, where the
CC collagen binding property and the HGF activity are maintained. The
CC present invention describes this HGF derivative as a live organ
CC transplant material that can be an artificial blood vessel or a stent in
CC which the collagen material is bound on its surface in a sheet-like
CC molding. Accordingly, it is useful for inducing and enhancing
CC angiogenesis (compared to natural type HGF) in damaged tissue and thus
CC can enhance tissue regeneration. Furthermore, it exhibits vulnerary
CC activity with increased stability and collagen binding properties. This
CC polynucleotide sequence is the DNA encoding the human fibronectin
CC collagen binding domain from Ala260 to Arg599, given in an
CC exemplification of the invention.

XX SQ Sequence 1051 BP; 274 A; 262 C; 281 G; 234 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.09e-09 Length: 1051
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-581-651D-9 (1-21) x ADS19042 (1-1051)

Qy 1 ValGlyAenGlyAArgGlyGluTrpThrCysIleAlaTySerGlnLeuArgAspGlnCys 20
Db 640 GTTGGGAATGTCGTGGGAATGGACATGCTACTTCGACGCTTCGAGATCAGTGC 699

Qy 21 Ile 21

Db 700 ATT 702

RESULT 2

AAAG4263
ID AAA64263 standard; DNA; 1053 BP.

XX AC AAA64263;

XX DT 20-DEC-2000 (first entry)

XX cDNA encoding a modified fibronectin collagen-binding domain.

XX Fibronectin; collagen-binding domain; sustained release; gene therapy;
KW physiologically active polypeptide; topical retention;
KW tissue regeneration; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers
FT CDS 16..1047
FT /tag= a
FT /product= "fibronectin collagen-binding domain"

FT WO200049159-A1.

XX PD 24-AUG-2000.

XX PF 21-FEB-2000; 2000WO-JP000964.

XX PR 19-FEB-1999; 99JP-00041913.

XX PR 01-NOV-1999; 99JP-00311364.

XX PA (TERU) TERUMO CORP.

XX

PI Ishikawa T, Kitajima T;

XX WPI; 2000-565375/52.

DR P-PSDB; AAB08505.

XX Collagen-binding active polypeptide for use in an agent for enabling

PT topical retention or sustained release of a physiologically active

PT peptide or physiological activity-imparting agent comprises a fibronectin

PT peptide.

XX Example 1; Page 126-127; 135pp; English.

XX The present sequence encodes a modified fibronectin collagen-binding
CC domain. The protein is used to construct a collagen-binding
CC physiologically active polypeptide. This polypeptide comprises a peptide
CC from fibronectin ligated to a physiologically active peptide. The
CC polypeptides are used in an agent for enabling topical retention or
CC sustained release of a physiologically active peptide or physiological
CC activity-imparting agent. They may be used in gene therapy and in tissue
CC regeneration

XX SQ Sequence 1053 BP; 275 A; 261 C; 281 G; 236 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.09e-09 Length: 1053
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-581-651D-9 (1-21) x AAA64263 (1-1053)

Qy 1 ValGlyAenGlyAArgGlyGluTrpThrCysIleAlaTySerGlnLeuArgAspGlnCys 20
Db 643 GTTGGGAATGTCGTGGGAATGGACATGCTACTTCGACGCTTCGAGATCAGTGC 702

Qy 21 Ile 21

Db 703 ATT 705

RESULT 3

ABL41020

ID ABL41020 standard; DNA; 1053 BP.

XX AC ABL41020;

XX DT 12-AUG-2002 (first entry)

XX Modified human fibronectin collagen-binding domain encoding DNA.

XX Osteogenesis; drug delivery system; DDS; collagen; osteopathic; human;
KW fibronectin; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX Key Location/Qualifiers
FT CDS 16..1047
FT /tag= a
FT /product= "modified fibronectin collagen-binding domain"

FT conflict 109

FT conflict 206

FT conflict 270

FT conflict 374

FT conflict 681

FT conflict 681

XX JP2002058485-A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 562.251 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-9
Perfect score: 120
Sequence: 1 VGNRGWETCIAYSLRDQCI 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=genEmbl -QFWT=fascap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -TRAN=human40.cdi -LIST=45
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CEN 1.1.9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sv:*
13: gb.un:*
14: gb.vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	100.0	358	6	CQ685590 Sequence
2	120	100.0	1053	6	BD133458 Collagen
3	120	100.0	1053	6	BD133469 Hybrid po
4	120	100.0	1053	6	E63263 Collagen-bi

5	120	100.0	1224	6	E63271	E63271 Collagen-bi
6	120	100.0	1407	6	BD133465	BD133465 Collagen-
7	120	100.0	1428	6	BD133476	BD133476 Hybrid po
8	120	100.0	1443	6	BD133466	BD133466 Collagen-
9	120	100.0	1527	6	E63270	E63270 Collagen-bi
10	120	100.0	1560	6	BD133477	BD133477 Hybrid po
11	120	100.0	1929	6	CQ871810	CQ871810 Sequence
12	120	100.0	1929	6	CQ871828	CQ871828 Sequence
13	120	100.0	2147	6	BD137021	BD137021 Polypteri
14	120	100.0	2147	6	AX003229	AX003229 Sequence
15	120	100.0	2147	9	HSA276395	HSA276395 Homo sapi
16	120	100.0	2192	9	HSA535086	HSA535086 Homo sapi
17	120	100.0	2358	6	CQ875358	CQ875358 Sequence
18	120	100.0	2409	6	CQ731571	CQ731571 Sequence
19	120	100.0	4321	9	HSM806214	HSM806214 Homo sapi
20	120	100.0	6816	6	CQ715726	CQ715726 Sequence
21	120	100.0	7299	9	HSM807162	HSM807162 Homo sapi
22	120	100.0	7323	11	BV178397	BV178397 sqm10026
23	120	100.0	7502	9	HSM804082	HSM804082 Homo sapi
24	120	100.0	7544	9	HSM803509	AL832202 Homo sapi
25	120	100.0	7679	6	AR274901	AR274901 Sequence
26	120	100.0	7680	6	AR380744	AR380744 Sequence
27	120	100.0	7680	6	AX277596	AX277596 Sequence
28	120	100.0	7680	6	AX335368	AX335368 Sequence
29	120	100.0	7680	9	HSPFBI	X02761 Human mRNA
30	120	100.0	7705	6	A14133	A14133 Fibronectin
31	120	100.0	7705	6	AR034630	AR034630 Sequence
32	120	100.0	7705	6	E01162	E01162 cDNA encodi
33	120	100.0	7705	6	I70110	I70110 Sequence 16
34	120	100.0	7705	6	AR364992	AR364992 Sequence
35	120	100.0	7753	9	AB191261	AB191261 Homo sapi
36	120	100.0	7803	6	AR051657	AR051657 Sequence
37	120	100.0	7868	9	HSM806653	BX640608 Homo sapi
38	120	100.0	7951	9	HSM806901	BX640802 Homo sapi
39	120	100.0	7951	9	HSM806902	BX649182 Homo sapi
40	120	100.0	8027	6	CQ833991	CQ833991 Sequence
41	120	100.0	8027	6	CQ875357	CQ875357 Sequence
42	120	100.0	8030	9	HSM806170	BX538017 Homo sapi
43	120	100.0	8035	9	HSM806171	BX538018 Homo sapi
44	120	100.0	8042	9	HSM806805	BX640731 Homo sapi
45	120	100.0	8044	6	AR454662	AR454662 Sequence

ALIGNMENTS

RESULT 1
CQ685590
LOCUS CQ685590 Sequence 30516 from Patent WO02070737.
DEFINITION CQ685590
ACCESSION CQ685590
VERSION CQ685590.1 GI:42213078
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liew, C.C., Marshall, W.B. and Zhang, H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 30516 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source
1. .358
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7.59e-10 Length: 358
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-581-651D-9 (1-21) x CQ685590 (1-358)
Qy 1 ValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 199 GTTGGGAATGTCGTGGGGAATGGACATGTCATTGCTACTCGACGCTTCGAGATCAGTGC 258
Qy 21 Ile 21
Db 259 ATT 261

RESULT 2
BD133458
LOCUS BD133458 1053 bp DNA linear PAT 18-SEP-2002
DEFINITION Collagen-binding osteogenesis promotion fused protein.
ACCESSION BD133458
VERSION BD133458.1 GI:23228403
KEYWORDS JP 2002058485-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
AUTHORS Kitajima,T. and Ishikawa,T.
TITLE Collagen-binding osteogenesis promotion fused protein
JOURNAL Patent: JP 2002058485-A 3 26-FEB-2002;
TERUMO CORP

COMMENT
OS Artificial Sequence
PN JP 2002058485-A/3
PD 26-FEB-2002

PI 16-AUG-2000 JP 200246744
PI TAKASHI KITAJIMA,TETSUYA ISHIKAWA
PC C12N15/09,A61K38/48,A61K48/00,A61P19/08,A61P19/10,C07K14/51,
PC C07K14/78,
PC

C07K19/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/00,A61K37/ PC
47,C12N5/00

CC Description of Artificial Sequence:Modified Human Fibronectin
CC Collagen-Binding Domain
FH Key Location/Qualifiers

FT conflict (109)
FT conflict (206)
FT conflict (270)
FT conflict (374)
FT conflict (681)
FT CDS (16)..(1044).
FEATURES
source
1..1053
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Alignment Scores:
Pred. No.: 2.33e-09 Length: 1053
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-9 (1-21) x BD133458 (1-1053)

Qy 1 ValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 643 GTTGGGAATGTCGTGGGGAATGGACATGTCATTGCTACTCGACGCTTCGAGATCAGTGC 702

Qy 21 Ile 21

Db 703 ATT 705

RESULT 3

BD133469
LOCUS BD133469 1053 bp DNA linear PAT 18-SEP-2002
DEFINITION Hybrid polypeptide having collagen-binding activity and neovascularization-adjusting activity.
ACCESSION BD133469
VERSION BD133469.1 GI:23228414
KEYWORDS JP 2002060400-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE
AUTHORS Ishikawa,T. and Kitajima,T.
TITLE Hybrid polypeptide having collagen-binding activity and neovascularization-adjusting activity
JOURNAL Patent: JP 2002060400-A 3 26-FEB-2002;
TERUMO CORP

COMMENT
OS Artificial Sequence
PN JP 2002060400-A/3
PD 26-FEB-2002

PI 17-AUG-2000 JP 2000247379
PI TETSUYA ISHIKAWA,TAKASHI KITAJIMA
PC C07K19/00,A61K38/00,A61K47/48,A61P9/10,C07K14/515,C07K14/78,
PC C12N1/15
PC C12N1/19,C12N1/21,C12N5/10//C12N15/09,A61K37/02,C12N5/00, PC
C12N15/00

CC Description of Artificial Sequence:Modified Human Fibronectin
CC Collagen-Binding Domain
FH Key Location/Qualifiers

FT conflict (109)
FT conflict (206)
FT conflict (270)
FT conflict (374)
FT conflict (681)
FT CDS (16)..(1044).
FEATURES
source
1..1053
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Alignment Scores:
Pred. No.: 2.33e-09 Length: 1053
Score: 120.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-9 (1-21) x BD133469 (1-1053)

Qy 1 ValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCys 20
Db 643 GTTGGGAATGTCGTGGGGAATGGACATGTCATTGCTACTCGACGCTTCGAGATCAGTGC 702

Qy 21 Ile 21

Db 703 ATT 705

RESULT 4

E63263
LOCUS E63263 1053 bp DNA linear PAT 27-AUG-2002
DEFINITION Collagen-binding physiologically active polypeptide.

ACCESSION E63263

VERSION E63263.1 GI:22556336

KEYWORDS JP 2001190280-A/3.

SOURCE synthetic construct

ORGANISM other sequences; artificial sequences.

REFERENCE 1 (bases 1 to 1053)

AUTHORS Ishikawa,T. and Kitajima,T.

TITLE Collagen-binding physiologically active polypeptide

JOURNAL Patent: JP 2001190280-A 3 17-JUL-2001;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 6.14379 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	749	2 Q7Z391	Q7Z391 homo sapien
2	110	100.0	1103	2 Q6MZFA	Q6MZFA homo sapien
3	110	100.0	2193	2 Q6MZM7	Q6MZM7 homo sapien
4	110	100.0	2240	2 Q68DP8	Q68DP8 homo sapien
5	110	100.0	2265	1 F1NC_BOVIN	P07589 bos taurus
6	110	100.0	2267	2 Q68DP9	Q68DP9 homo sapien
7	110	100.0	2296	2 Q6N0A6	Q6N0A6 homo sapien
8	110	100.0	2357	2 Q68DT4	Q68DT4 homo sapien
9	110	100.0	2444	2 Q6N025	Q6N025 homo sapien
10	110	100.0	2477	2 Q6MZU5	Q6MZU5 homo sapien
11	106	96.4	2386	1 F1NC_HUMAN	P02751 homo sapien
12	105	95.5	2477	1 F1NC_MOUSE	P11276 mus musculus
13	105	95.5	2477	1 F1NC_RAT	P04937 rattus norv
14	81	73.6	2481	1 F1NC_XENLA	Q91740 xenopus lae
15	81	73.6	2481	2 Q6GOA5	Q6GOA5 xenopus lae
16	75	68.2	922	2 Q61405	O33405 brachydanio
17	75	68.2	2478	2 Q63406	O93406 brachydanio
18	73	66.4	2408	2 Q6JAN2	Q6JAN2 brachydanio
19	57	51.8	670	2 Q6DF16	Q6DF16 xenopus tro
20	57	51.8	674	2 Q68TC6	Q68TC6 cyprinus ca
21	57	51.8	680	2 Q7T317	Q7T317 brachydanio
22	56	50.9	686	2 Q6DE15	Q6DE15 gallus gall
23	55	50.0	707	1 MM09_HUMAN	P14780 homo sapien
24	54	49.1	679	2 Q68856	Q68856 cynops pyrr
25	53	48.2	730	1 MM09_MOUSE	P41245 mus musculus
26	53	48.2	730	2 Q60XI8	Q60XI8 mus musculus
27	52	47.3	3519	1 OL56_STRAT	Q07017 streptomyc
28	52	47.3	3150	2 Q6KIV4	Q6KIV4 streptomyc
29	51	46.4	921	2 Q6BUC1	Q6BUC1 mus musculus
30	51	46.4	1307	2 Q60V33	Q60V33 mus musculus
31	51	46.4	2078	2 Q69ZN7	Q69ZN7 mus musculus

32 50 45.5 212 2 Q9T1D8 Q9T1D8 bacterioph
33 50 45.5 675 2 Q8QFQ6 Q8QFQ6 oncorhynch
34 49 44.5 94 2 Q86MY2 Q86MY2 rhodnius pr
35 49 44.5 458 2 Q6Y7P9 Q6Y7P9 staphylococ
36 49 44.5 671 2 Q6PE33 Q6PE33 xenopus lae
37 49 44.5 671 2 Q9W7L6 Q9W7L6 xenopus lae
38 49 44.5 690 2 Q9PVN5 Q9PVN5 oryzias lat
39 48 43.6 94 2 Q86MY3 Q86MY3 rhodnius pr
40 48 43.6 208 2 Q9XSD1 Q9XSD1 canis famill
41 48 43.6 673 2 Q90VB3 Q90VB3 paralicthy
42 48 43.6 704 1 MM09_CANFA O18733 canis famill
43 48 43.6 704 2 Q71U09 O18733 canis famill
44 48 43.6 747 2 Q9S241 Q9S241 streptomyc
45 47 42.7 157 2 P94805 P94805 haloterax a

ALIGNMENTS

RESULT 1

ID Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Blocher H., Boecher M., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 110; DB 2; Length 749;

Best Local Similarity 100.0%; Pred. No. 6.3e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||
Db 494 TDHTVLVQTRGNSNGALCH 513

RESULT 2

Q6MZFA PRELIMINARY; PRT; 1103 AA.
ID Q6MZFA
AC Q6MZFA;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219)

DE (Fragment).

```

GN Name=DKFZp686K139; Synonym=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansoerge W., Krieger S., Regliet T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE46200.1; -.
DR GO; BX640802; CAE45885.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00041; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS0853; FN3; 4.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;

Query Match 100.0%; Score 110; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGGNSGALCH 20
DB 491 TDHTVLVQTRGGNSGALCH 510

RESULT 3
OSMZW7
AC Q6MZW7 PRELIMINARY; PRT; 2193 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein DKFZp686O12165 (Fragment).
GN Name=DKFZp686O12165;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG The German Human cDNA Consortium;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640999; CAE46002.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR02086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003962; FN_III subd.
DR InterPro; IPR003961; FN_III.

Name=DKFZp686K139; Synonym=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Human cDNA Consortium;
RG Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003962; FN_III subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.

```

```

DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 7.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 7.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 17.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 7.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS0853; FN3; 17.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 2193 AA; 240641 MW; F876B93106540EF3 CRC64;

Query Match 100.0%; Score 110; DB 2; Length 2193;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGGNSGALCH 20
DB 149 TDHTVLVQTRGGNSGALCH 168

RESULT 4
Q68DP8
AC Q68DP8 PRELIMINARY; PRT; 2240 AA.
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein DKFZp686H0342.
GN Name=DKFZp686H0342;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003962; FN_III subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.51634 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79;*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	2265	1 FNBO	fibronectin - bovi
2	106	96.4	2386	1 FNHU	fibronectin precu
3	105	95.5	2477	2 S14428	fibronectin precu
4	81	73.6	2481	2 A43908	fibronectin - Afri.
5	55	50.0	707	1 A34458	gelatinase B (EC 3
6	53	48.2	730	1 I52580	gelatinase B (EC 3
7	53	48.2	730	2 J01456	gelatinase B (EC 3
8	52	47.3	3519	2 S43048	polyketide synthas
9	48	43.6	747	2 T36812	probable dehydroge
10	47	42.7	157	2 T44794	hypothetical prote
11	47	42.7	340	2 T32931	hypothetical prote
12	47	42.7	340	2 C87732	protein W03D8.4 [i
13	47	42.7	438	1 A57667	pop-1 protein - Ca
14	47	42.7	708	2 J04364	gelatinase B (EC 3
15	47	42.7	708	2 S62907	gelatinase B (EC 3
16	47	42.7	1079	2 A70972	probable DNA polym
17	46	41.8	493	2 T06031	hexokinase homolog
18	46	41.8	707	1 A53796	gelatinase B (EC 3
19	46	41.8	712	1 I46031	gelatinase B (EC 3
20	44	40.0	413	2 J02135	chitinase (EC 3.2.
21	44	40.0	606	2 T09892	hypothetical prote
22	44	40.0	843	2 T06068	probable proton pu
23	43	39.1	237	2 A98357	hypothetical prote
24	43	39.1	237	2 A92925	hypothetical prote
25	43	39.1	692	2 H69961	transcription regu
26	43	39.1	1430	2 T12449	hypothetical prote
27	42.5	38.6	183	2 D87614	hypothetical prote
28	42	38.2	288	2 T10477	sec13 protein - ye
29	42	38.2	801	2 H83737	glucosidase BH0704

RESULT 1

FNBO

fibronectin - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A26452; B21165; A23292

R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A:Title: Complete primary structure of bovine plasma fibronectin.

A:Reference number: A26452; MUID:87054047; PMID:3780752

A:Accession: A26452

A:Molecule type: protein

A:Residues: 1-2265 <SKO>

A:Cross-references: UNIPROT:P07589

R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983

A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

A:Reference number: A21165; MUID:83221567; PMID:6304699

A:Accession: B21165

A:Molecule type: mRNA

A:Residues: 2170-2265 <KOR>

A:Cross-references: GB:K00800; NID:gl63055; PIDN:AAA30521.2; PID:95713333

R:Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sott

Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983

A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal

A:Reference number: A23292; MUID:83117805; PMID:6218503

A:Accession: A23292

A:Molecule type: protein

A:Residues: 1-16, 'C', 18-20, 'S', 22-432:447-463;1367-1517;1567-1673;2062-2176, 'N', 2178-226

C:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.

C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b

C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib

aling, and maintenance of cell shape.

C:Comment: Plasma fibronectin is synthesized by hepatocytes.

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep

C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu

R:21-241/Domain: fibrin and heparin binding <FBR>

F:21-56/Domain: fibronectin type I repeat homology <1F1>

F:66-104/Domain: fibronectin type I repeat homology <1F2>

F:110-148/Domain: fibronectin type I repeat homology <1F3>

F:155-194/Domain: fibronectin type I repeat homology <1F4>

F:200-239/Domain: fibronectin type I repeat homology <1F5>

F:277-577/Domain: collagen binding <CB>

F:277-311/Domain: fibronectin type I repeat homology <1F6>

F:329-370/Domain: fibronectin type II repeat homology <2F1>

F:389-430/Domain: fibronectin type II repeat homology <2F2>

F:439-477/Domain: fibronectin type I repeat homology <1F7>

F:487-524/Domain: fibronectin type I repeat homology <1F8>

F:530-568/Domain: fibronectin type I repeat homology <1F9>

F:578-661/Domain: fibronectin type III repeat homology <FN3A>

F:688-770/Domain: fibronectin type III repeat homology <FN3B>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

unknown protein P5
BLN protein [limo
vpr protein - huma
gas-vesicle operon
gas-vesicle operon
fibronectin - east
tegatose-1,6-bisph
cell division prot
hypothetical prote
gelatinase A (EC 3
gelatinase A (EC 3
gelatinase A (EC 3
gelatinase A (EC 3
gelatinase A (EC 3
hypothetical prote
ribonucleoside-dip

F:875-957/Domain: fibronectin type III repeat homology <FN3D>
 F:965-1046/Domain: fibronectin type III repeat homology <FN3E>
 F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
 F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
 F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>
 F:1326-1404/Domain: fibronectin type III repeat homology <FN3I>
 F:1410-1517/Domain: cell attachment <CAD>
 F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
 F:1431-1495/Region: cell attachment (R-G-D) motif
 F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
 F:1600-1870/Domain: heparin binding <HB2>
 F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>
 F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>
 F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>
 F:1970-1972/Region: cell attachment (R-G-D) motif
 F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
 F:1985-2216/Domain: fibrin binding <FB>
 F:2085-2124/Domain: fibronectin type I repeat homology <IF10>
 F:2130-2167/Domain: fibronectin type I repeat homology <IF11>
 F:2174-2209/Domain: fibronectin type I repeat homology <IF12>
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F:3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status expirime
 F:21-47 45-56 66-94 92-104 110-138 136-148 155-184 182-194 200-229 227-239 277-304 302-3
 7 2155-2167 2174-2200 2198-2209/Disulfide bonds: #status predicted
 F:399 497 511 846 976 1213 1987/Binding site: carbonylhydrate (Asn) (covalent) #status expe
 F:1205 1692/Binding site: carbonylhydrate (Asn) (covalent) #status absent
 F:1943 1944/Binding site: carbonylhydrate (Thr) (covalent) #status experimental
 F:2246/Disulfide bonds: interchain (to 2250) #status predicted
 F:2250/Disulfide bonds: interchain (to 2246) #status predicted
 F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 100.0%; Score 110; DB 1; Length 2265;
 Best Local Similarity 100.0%; Pred. NO. 2.1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTLVQTRGNSGNALCH 20
 DB 371 TDHTLVQTRGNSGNALCH 390

RESULT 2

FNHU
 fibronectin precursor [validated] - human
 N:Alternate names: fibronectin splice form ED-A
 C:Species: Homo sapiens (man)
 C:Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004
 A:Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
 R:Dean, D.C.; Bowls, C.L.; Bourgeois, S.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
 A:Title: Cloning and analysis of the promoter region of the human fibronectin gene.
 A:Reference number: A26460; MUID:87175578; PMID:3031656
 A:Accession: A26460
 A:Molecule type: DNA
 A:Residues: 1-49 <DEA>
 A:Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:G182686; PIDN:AAAS337
 R:Oldberg, A.; Ruoslahti, E.
 J. Biol. Chem. 261, 2113-2116, 1986
 A:Title: Evolution of the fibronectin gene.
 A:Reference number: A26284; MUID:86111901; PMID:3003095
 A:Accession: A26284
 A:Molecule type: DNA
 A:Residues: 1447-1540 <OLD>
 A:Cross-references: GB:M12549; NID:G182688
 A>Note: The authors translated the codon TTC for residue 1494 as Glu
 R:Paolella, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
 Nucleic Acids Res. 16, 3545-3557, 1988
 A:Title: Sequence analysis and in vivo expression show that alternative splicing of ED-E
 A:Reference number: S00848; MUID:882333940; PMID:3375063
 A:Accession: S03917
 A:Molecule type: DNA
 A:Residues: 1594-1767 'V', 1769-1783 <PAO>
 A:Cross-references: EMBL:X07718; NID:G31402
 A>Note: the authors translated the codon AAC for residue 1631 as Asp

R:Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
 FEBS Lett. 207, 287-291, 1986
 A:Title: Donor and acceptor splice signals within an exon of the human fibronectin gene.
 A:Reference number: A24854; MUID:87030929; PMID:3770201
 A:Accession: A24854
 A:Molecule type: DNA
 A:Residues: 1992-2147 <VIB>
 A:Cross-references: GB:X04530; NID:G31436
 R:Gutman, A.; Yamada, K.M.; Kornblihtt, A.
 FEBS Lett. 207, 145-148, 1986
 A:Title: Human fibronectin is synthesized as a pre-propolypeptide.
 A:Reference number: A24476; MUID:87030890; PMID:3770189
 A:Accession: A24476
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-14, 'Q', 16-38 <GUT>
 R:Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
 EMBO J. 4, 1755-1759, 1985
 A:Title: Primary structure of human fibronectin: differential splicing may generate at l
 A:Reference number: A91008; MUID:85284965; PMID:2992939
 A:Accession: A91008
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
 A:Cross-references: GB:X02761
 R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Nucleic Acids Res. 12, 5853-5868, 1984
 A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
 A:Reference number: A93529; MUID:84272258; PMID:6462919
 A:Accession: A93529
 A:Molecule type: mRNA
 A:Residues: 973-2080; 2112-2386 <KO2>
 A:Cross-references: GB:X00739
 R:Oldberg, A.; Linney, E.; Ruoslahti, E.
 J. Biol. Chem. 258, 10193-10196, 1983
 A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
 A:Reference number: A21011; MUID:83290929; PMID:6688418
 A:Accession: A21011
 A:Molecule type: mRNA
 A:Residues: 1434-1537 <OL2>
 A:Cross-references: GB:X00055; NID:G182680; PIDN:AAAS2459.1; PID:G182683
 R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
 Biochemistry 24, 2698-2704, 1985
 A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with
 A:Reference number: A90495; MUID:85280409; PMID:2992573
 A:Accession: A90495
 A:Molecule type: mRNA
 A:Residues: 1594-2386 <BER>
 A:Cross-references: GB:M10905; NID:G182696; PIDN:AAAS2462.1; PID:G182697
 R:Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.
 FEBS Lett. 186, 31-34, 1985
 A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
 A:Reference number: A22245; MUID:85231203; PMID:2989004
 A:Accession: A22245
 A:Molecule type: mRNA
 A:Residues: 1948-2067 <UME>
 A:Cross-references: GB:M27589; NID:G182705; PIDN:AAAS2465.1; PID:G182706
 A:Accession: B22245
 A:Molecule type: mRNA
 A:Residues: 1975-1991; 2017-2039 <UM2>
 A:Cross-references: GB:M27590
 R:Seiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
 Biochemistry 25, 4936-4941, 1986
 A:Title: Human liver fibronectin complementary DNAs: identification of two different mes
 A:Reference number: 152394; MUID:87026578; PMID:3021206
 A:Accession: I65273
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
 A:Cross-references: GB:M14060; NID:G182701; PIDN:AAAS2464.1; PID:G182704
 R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 6.77124 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	110	100.0	228	US-10-344-634-4
2	110	100.0	285	US-10-344-634-16
3	110	100.0	343	US-09-934-706-1
4	110	100.0	386	US-10-344-634-14
5	110	100.0	400	US-09-934-706-5
6	110	100.0	501	US-09-934-706-4
7	110	100.0	642	US-10-741-601-354
8	110	100.0	642	US-10-741-601-1066
9	110	100.0	657	US-10-741-601-359
10	110	100.0	657	US-10-741-600-1072
11	110	100.0	984	US-10-741-601-356

12	110	100.0	984	US-10-741-600-1069	Sequence 1069, Ap
13	110	100.0	1173	US-10-450-763-52634	Sequence 52634, A
14	110	100.0	2220	US-10-236-392-4	Sequence 4, Appli
15	110	100.0	2296	US-10-741-601-363	Sequence 363, App
16	110	100.0	2296	US-10-741-600-1075	Sequence 1075, Ap
17	110	100.0	2320	US-10-279-733-8	Sequence 8, Appli
18	110	100.0	2355	US-10-144-194A-104	Sequence 104, App
19	110	100.0	2355	US-10-447-161-3	Sequence 3, Appli
20	110	100.0	2355	US-10-734-564-94	Sequence 564, Appl
21	110	100.0	2355	US-10-741-601-357	Sequence 357, App
22	110	100.0	2355	US-10-741-601-366	Sequence 366, App
23	110	100.0	2355	US-10-491-566-104	Sequence 104, App
24	110	100.0	2355	US-10-741-600-1067	Sequence 1067, Ap
25	110	100.0	2355	US-10-741-600-1078	Sequence 1078, App
26	110	100.0	2355	US-10-852-335A-147	Sequence 147, App
27	110	100.0	2355	US-10-287-436A-436	Sequence 436, App
28	110	100.0	2355	US-10-287-436A-1137	Sequence 1137, Ap
29	110	100.0	2355	US-11-040-130-28	Sequence 28, Appl
30	110	100.0	2386	US-10-741-601-360	Sequence 360, App
31	110	100.0	2386	US-10-618-281-32	Sequence 1071, Ap
32	110	100.0	2386	US-10-741-600-1071	Sequence 2, Appli
33	106	96.4	2320	US-10-236-392-2	Sequence 64, Appl
34	106	96.4	2328	US-10-171-311-64	Sequence 70, Appl
35	106	96.4	2328	US-10-236-031B-70	Sequence 98, Appl
36	106	96.4	2328	US-10-374-979-98	Sequence 98, Appl
37	106	96.4	2328	US-10-182-936A-98	Sequence 98, Appl
38	106	96.4	2328	US-10-477-238A-677	Sequence 677, App
39	106	96.4	2328	US-10-680-287A-677	Sequence 677, App
40	106	96.4	2328	US-10-477-173-677	Sequence 677, App
41	106	96.4	2355	US-10-360-101-235	Sequence 235, App
42	106	96.4	2386	US-09-961-403-1	Sequence 1, Appli
43	106	96.4	2386	US-10-788-792-206	Sequence 206, App
44	106	96.4	2386	US-10-868-577A-59	Sequence 59, Appl
45	106	96.4	2386	US-10-485-758-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-344-634-4
; Sequence 4, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takaishi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human Fibronectin
; OTHER INFORMATION: Collagen-binding Domain
US-10-344-634-4

Query Match 100.0%; Score 110; DB 15; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTRGGNSNGALCH 20
Db 113 TDHTVLVQTRGGNSNGALCH 132

RESULT 2
US-10-344-634-16
; Sequence 16, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
; OTHER INFORMATION: Epidermal Growth Factor
US-10-344-634-16
Query Match 100.0%; Score 110; DB 15; Length 285;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||
Db 113 TDHTVLVQTRGNSNGALCH 132
|||||
RESULT 3
US-09-934-706-1
; Sequence 1, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified Human
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: CONFLICT
; LOCATION: (69)
; NAME/KEY: CONFLICT
; LOCATION: (125)
US-09-934-706-1
Query Match 100.0%; Score 110; DB 9; Length 343;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||

Db 113 TDHTVLVQTRGNSNGALCH 132
|||||
RESULT 4
US-10-344-634-14
; Sequence 14, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
; OTHER INFORMATION: Basic Fibroblast Growth Factor
US-10-344-634-14
Query Match 100.0%; Score 110; DB 15; Length 386;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TDHTVLVQTRGNSNGALCH 20
|||||
Db 113 TDHTVLVQTRGNSNGALCH 132
|||||
RESULT 5
US-09-934-706-5
; Sequence 5, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 5
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Hybrid
; OTHER INFORMATION: Polypeptide of Human Fibronectin Collagen-Binding
; OTHER INFORMATION: Domain and Human Epidermal Growth Factor
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)..
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: PEPTIDE
; LOCATION: (343)..
; OTHER INFORMATION: /note=" enterokinase recognition sequence"
; NAME/KEY: PEPTIDE
; LOCATION: (348)..
; OTHER INFORMATION: /note="human epidermal growth factor"
US-09-934-706-5
Query Match 100.0%; Score 110; DB 9; Length 400;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.11765 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-8
Perfect score: 110
Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	96.4	2231	1	US-08-153-799-16
2	106	96.4	2324	1	US-08-283-857-1
3	106	96.4	2324	5	PCT-US95-09819-1
4	106	96.4	2327	6	5455158-1
5	106	96.4	2327	6	5455158-1
6	106	96.4	2386	2	US-09-016-366A-12
7	106	96.4	2386	4	US-09-961-403-1
8	106	96.4	2446	2	US-08-551-356-2
9	106	96.4	2446	5	PCT-US93-12687-2
10	87	79.1	65	2	US-08-982-597A-19
11	87	79.1	65	3	US-09-136-218-19
12	55	50.0	707	3	US-08-704-711A-19
13	55	50.0	707	3	US-09-521-220-19
14	55	50.0	707	3	US-09-391-104-20
15	55	50.0	707	4	US-09-949-016-6575
16	55	50.0	708	3	US-08-448-489-16
17	55	50.0	708	4	US-09-689-730-16
18	55	50.0	713	4	US-09-949-016-10629
19	53	48.2	60	2	US-08-982-597A-18
20	53	48.2	60	3	US-09-136-218-18
21	52	47.3	3519	3	US-09-428-517-4
22	52	47.3	4150	3	US-09-428-517-2
23	48	43.6	13	2	US-08-982-597A-8
24	48	43.6	13	3	US-09-136-218-8
25	48	43.6	14	2	US-08-982-597A-9
26	48	43.6	14	3	US-09-136-218-9
27	48	43.6	15	2	US-08-982-597A-1

Sequence 1, Appli
Sequence 33829, A
Sequence 49046, A
Sequence 14, Appl
Sequence 14, Appl
Sequence 44610, A
Sequence 1, Appli
Sequence 6520, Ap
Sequence 8335, Ap
Sequence 7, Appli
Sequence 15, Appl
Sequence 15, Appl
Sequence 23, Appl
Sequence 23, Appl
Sequence 59273, A
Sequence 374, App
Sequence 4801, Ap

ALIGNMENTS

RESULT 1
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: Protein
/ LOCATION: 1..2231
/ OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match          96.4%; Score 106; DB 1; Length 2321;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTGGNSGALCH 20
Db 371 TDHTVLVQTGGNSGALCH 390

RESULT 2
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5792742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283,857
; FILING DATE: 01-AUG-1994
; PRIOR APPLICATION NUMBER: US 07/714,134
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-09819-1

Query Match          96.4%; Score 106; DB 5; Length 2324;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTGGNSGALCH 20
Db 371 TDHTVLVQTGGNSGALCH 390

RESULT 3
PCT-US95-09819-1
; Sequence 1, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-09819-1

Query Match          96.4%; Score 106; DB 5; Length 2324;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TDHTVLVQTGGNSGALCH 20
Db 371 TDHTVLVQTGGNSGALCH 390

RESULT 4
5455158-1
; Patent No. 5455158
; APPLICANT: VOGEL, TIKVA; LEVANON, AVIGDOR; WERBER, MOSHE M.;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.39869 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	20	2	AAY28913
2	110	100.0	21	2	AAY28905
3	110	100.0	228	5	ABB06747 Human fib
4	110	100.0	285	5	ABB06751 Collagen-
5	110	100.0	343	3	ABB08505 Amino aci
6	110	100.0	343	5	ABB07961 Modified
7	110	100.0	386	5	ABB06750 Collagen-
8	110	100.0	400	3	ABB08509 Hybrid of
9	110	100.0	461	5	ABB07964 Human fib
10	110	100.0	473	5	ABB07965 Human fib
11	110	100.0	501	3	ABB08508 Hybrid of
12	110	100.0	642	8	ADQ39403 Human myo
13	110	100.0	642	8	ADR67316 Human bla
14	110	100.0	642	8	ADP17489 Amino aci
15	110	100.0	642	8	ADR97658 Human fib
16	110	100.0	657	8	ADQ39409 Human myo
17	110	100.0	660	2	AAY28901 Human mig
18	110	100.0	951	8	ADP19044 Chimeric
19	110	100.0	984	8	ADQ39406 Human myo
20	110	100.0	1173	4	ABG22275 Novel hum
21	110	100.0	1179	8	ADP75952 Human min
22	110	100.0	1223	8	ADP30365 Human sec
23	110	100.0	1359	8	ADP75957 Human leu
24	110	100.0	2182	8	ADR66462 Human pro
25	110	100.0	2182	8	ADR66120 Human pro

26	110	100.0	2220	6	ABO01289 Human pro
27	110	100.0	2265	4	AAM38647 Human pol
28	110	100.0	2265	8	ADS17498 Amino aci
29	110	100.0	2265	8	ADR97667 Bovine fi
30	110	100.0	2265	6	ABR40124 Human cel
31	110	100.0	2296	8	ADQ39412 Human myo
32	110	100.0	2320	6	AAE37107 Human fib
33	110	100.0	2320	6	ABR42588 Human fib
34	110	100.0	2330	4	AAM38646 Human pol
35	110	100.0	2355	4	AAM38649 Human pol
36	110	100.0	2355	6	ABR58335 NM_00202
37	110	100.0	2355	7	ADP65196 Human fib
38	110	100.0	2355	8	ADG89560 Human fib
39	110	100.0	2355	8	ADO55175 Protein #
40	110	100.0	2355	8	ADQ26085 Fibronect
41	110	100.0	2355	8	ADQ29668 Human col
42	110	100.0	2355	8	ADQ39415 Human myo
43	110	100.0	2355	8	ADQ39404 Human myo
44	110	100.0	2355	8	ADR67315 Human bla
45	110	100.0	2386	4	AAM38648 Human pol

ALIGNMENTS

RESULT 1

AAY28913
ID AAY28913 standard; peptide; 20 AA.

AC AAY28913;

DT 21-SEP-1999 (first entry)

DE MSF 1-alpha peptide epitope.

KW Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.

OS Synthetic.

OS Homo sapiens.

PN WO9931233-A1.

XX 24-JUN-1999.

PF 15-DEC-1998; 98WO-GB003766.

PR 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI, 1999-430039/36.

XX Proteins with cell migration stimulatory activity used in treating wound

XX and preventing scarring.

XX Claim 29; Page 61; 86pp; English.

CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28912-913 represent peptide epitopes of MSF against which monoclonal antibodies can be raised

SQ Sequence 20 AA;

Query Match 100.0%; Score 110; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 5.3e-11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGNSNGALCH 20
DB 1 TDHTVLVQTRGNSNGALCH 20

RESULT 2
AAAY28905
ID AAY28905 standard; peptide; 21 AA.
XX
AC AAY28905;
XX
DT 21-SEP-1999 (first entry)
XX
DE MSF 1-alpha peptide epitope.
XX
KW Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.
XX
OS Synthetic.
OS Homo sapiens.
PN WO9931233-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-GB003766.
XX
PR 16-DEC-1997; 97GB-00026539.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Schor SL, Schor AM;
XX
DR WPI; 1999-430039/36.
XX
PT Proteins with cell migration stimulatory activity used in treating wound
PT and preventing scarring.
XX
PS Claim 18; Page 57; 86pp; English.
XX
CC The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. Sequences AAY28902-906
CC represent peptide epitopes of MSF against which monoclonal antibodies
CC that are specific to MSF and do not cross-react with fibronectin are
CC raised
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 110; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.6e-11; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGNSNGALCH 20
DB 2 TDHTVLVQTRGNSNGALCH 21

RESULT 3
ABB06747
ID ABB06747 standard; protein; 228 AA.
XX
AC ABB06747;
XX
DT 13-JUN-2002 (first entry)
XX
DE Human fibronectin collagen-binding domain protein SEQ ID NO:4.
XX
KW Human; fibronectin; collagen-binding domain; collagen-binding hybrid;
KW basic fibroblast growth factor; enterokinase recognition site; collagen;
KW basic fibroblast growth factor; enterokinase recognition site; collagen;

epidermal growth factor; drug delivery system; tissue regeneration.
XX
OS Homo sapiens.
PN WO200214505-A1.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-JP007036.
XX
PR 15-AUG-2000; 2000JP-00246341.
XX
PA (TERU) TERUMO CORP.
XX
PI Ishikawa T, Kitajima T;
XX
DR WPI; 2002-257605/30.
DR N-PSDB; ABL50261.
XX
PT Collagen-binding hybrid polypeptide, useful in drug delivery system for
PT functional polypeptides, formulating into complex with collagen to give
PT functionally modified collagen matrix as biomaterial for tissue
PT regeneration.
XX
PS Example 1; Page 61-63; 86pp; Japanese.
XX
CC The present invention describes a hybrid polypeptide comprising a
CC collagen-binding domain composed of an amino acid sequence ranging from
CC Ala at position 260 to Arg at position 484 of human fibronectin or a
CC similar amino acid sequence but with some amino acids deleted,
CC substituted, inserted or added, and a functional polypeptide linked
CC together. The present invention also describes: (1) a biomaterial
CC containing a functional polypeptide-modified collagen obtained by
CC formulating a polypeptide originated from the hybrid polypeptide with
CC collagen into a complex; (2) a gene encoding the hybrid polypeptide; and
CC (3) a transformant containing the gene. The hybrid polypeptide is useful
CC in drug delivery system for functional polypeptides, formulating into
CC complex with collagen to give functionally modified collagen matrix as
CC biomaterial for tissue regeneration. The hybrid polypeptide has superior
CC activity and stability in the body for a long period of time, with
CC localisability and sustained-releasability. The present sequence
CC represents the human fibronectin collagen-binding domain, which is used
CC in an example from the present invention
XX
SQ Sequence 228 AA;
Query Match 100.0%; Score 110; DB 5; Length 228;
Best Local Similarity 100.0%; Pred. No. 9.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDHTVLVQTRGNSNGALCH 20
DB 113 TDHTVLVQTRGNSNGALCH 132

RESULT 4
ABB06751
ID ABB06751 standard; protein; 285 AA.
XX
AC ABB06751;
XX
DT 13-JUN-2002 (first entry)
XX
DE Collagen-binding epidermal growth factor protein SEQ ID NO:16.
XX
KW Human; fibronectin; collagen-binding domain; collagen-binding hybrid;
KW basic fibroblast growth factor; enterokinase recognition site; collagen;
KW epidermal growth factor; drug delivery system; tissue regeneration.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200214505-A1.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 415.948 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSGALCH 20

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q/cgm2_1/USPTO-spool_h/US9581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9581651 @CGN 1 1 8190 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	110	100.0	263	2	BE699322 RC3-NN006
2	110	100.0	310	1	AA375384 EST87654
3	110	100.0	373	1	AA376281 EST88736
4	110	100.0	385	1	AL589418 DKF2p451P
C 5	110	100.0	392	2	BE079849 RC6-BT062
C 6	110	100.0	400	2	BE122858 14_08 Hum
7	110	100.0	411	5	BX642740 DKF2p781G
8	110	100.0	458	1	AJ661698 AJ661698
9	110	100.0	471	4	BG007031 RC1-GN023

10	110	100.0	471	5	BX495442
11	110	100.0	478	5	BX495442
12	110	100.0	480	7	CN419466
13	110	100.0	484	4	BF986449
14	110	100.0	528	7	CF793207
15	110	100.0	538	1	AJ685866
C 16	110	100.0	542	5	BQ340573
17	110	100.0	558	7	CN419467
18	110	100.0	560	4	BG897886
19	110	100.0	605	7	CN481619
20	110	100.0	618	7	CN482377
21	110	100.0	620	7	CN481412
22	110	100.0	622	7	CN419625
23	110	100.0	632	7	CN482351
24	110	100.0	645	7	CN419499
25	110	100.0	653	7	CN419474
26	110	100.0	655	7	CN419546
27	110	100.0	657	7	CN419503
28	110	100.0	664	6	CB471201
29	110	100.0	673	1	AU140450
30	110	100.0	695	7	CN419477
31	110	100.0	702	7	CN419472
32	110	100.0	705	1	AU140899
33	110	100.0	723	7	CN419622
C 34	110	100.0	725	5	BQ574857
35	110	100.0	730	1	AU140834
36	110	100.0	731	1	AU140526
37	110	100.0	748	1	AU140948
38	110	100.0	748	1	AU140991
39	110	100.0	758	1	AU140556
40	110	100.0	766	7	CN419594
C 41	110	100.0	780	7	CN162843
42	110	100.0	781	7	CN161442
43	110	100.0	796	1	AU140814
44	110	100.0	816	1	AU141008
C 45	110	100.0	822	1	AA788933

ALIGNMENTS

RESULT 1
BE699322/c

LOCUS
RC3-NN0064-150600-022-all

ACCESSION
BE699322

VERSION
BE699322.1

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 263)

AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brundsen, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
20202663

PUBMED
10737800

COMMENT
Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

BE699322 263 bp mRNA linear EST 12-SEP-2000
RC3-NN0064-150600-022-all NN0064 Homo sapiens CDNA, mRNA sequence.

BE699322 263 bp mRNA linear EST 12-SEP-2000
RC3-NN0064-150600-022-all NN0064 Homo sapiens CDNA, mRNA sequence.

BE699322.1 GI:10087064

EST.

Homo sapiens (human)

Homo sapiens

1 (bases 1 to 263)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brundsen, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=dt2-RC3-NN0064-150
600-022-all&t3=2000-06-15&t4=1)

Seq primer: puc 18 forward
High quality sequence stop: 263.

Location/Qualifiers

FEATURES

source

```
1..263
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN0064"
/note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,466-08 Length: 263
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
```

US-09-581-651D-8 (1-20) x BE699322 (1-263)

```
QY 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCysHis 20
|||||
Db 129 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCGAATGGTGGCTTGTGCCAC 70
```

RESULT 2

```
AA375384 310 bp mRNA linear EST 21-APR-1997
LOCUS EST87654 HSC172 cells II Homo sapiens cDNA 5' end similar to
DEFINITION fibronectin, mRNA sequence.
```

ACCESSION AA375384

VERSION AA375384.1 GI:2027703

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

ADAMS, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns

Nature 377 (6347 Suppl), 3-174 (1995)

JOURNAL

96026280

MEDLINE

7566098

PUBMED

Other ESTs: THC167580

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tadb/hgi/hgi.html>)

Seq primer: M13 Reverse

Location/Qualifiers

1..310

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):179810"

/db_xref="taxon:9606"

/cell_type="fibroblast"

/cell_line="HSC172 (60PDL)"

/dev_stage="fetal"

/clone_lib="HSC172 cells II"

/note="Organ: lung; Vector: pBluescript SK-; Site_1:

ECORI; Site_2: XhoI"

ORIGIN

```
Alignment Scores:
Pred. No.: 2,98e-08 Length: 310
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
```

US-09-581-651D-8 (1-20) x AA375384 (1-310)

```
QY 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCysHis 20
|||||
Db 19 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCGAATGGTGGCTTGTGCCAC 78
```

RESULT 3

AA376281

LOCUS

DEFINITION

EST88736 HSC172 cells II Homo sapiens cDNA 5' end similar to

fibronectin, mRNA sequence.

ACCESSION AA376281

VERSION AA376281.1 GI:2028652

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 373)

ADAMS, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meisner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns

Nature 377 (6347 Suppl), 3-174 (1995)

JOURNAL

96026280

MEDLINE

7566098

PUBMED

Other ESTs: THC167580

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 102.693 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-8
Perfect score: 110
Sequence: 1 TDHTVLVQTRGNSNGALCH 20
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=PublishedApplications_NA -QPMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALICN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09581651 @cgn1_1.1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	310	9	US-09-765-231A-1
2	110	100.0	708	19	US-10-344-634-3
3	110	100.0	731	15	US-10-198-846-8918
4	110	100.0	879	19	US-10-344-634-15
5	110	100.0	1053	9	US-09-934-706-8
6	110	100.0	1182	19	US-10-344-634-13
7	110	100.0	1189	15	US-10-198-846-11556
8	110	100.0	1224	9	US-09-934-706-16
9	110	100.0	1527	9	US-09-934-706-15
10	110	100.0	2127	17	US-10-210-120-49
11	110	100.0	2127	22	US-10-956-157-4288
12	110	100.0	2127	24	US-10-909-035-49
13	110	100.0	2443	20	US-10-741-601-70
14	110	100.0	2443	22	US-10-741-600-238
15	110	100.0	2488	20	US-10-741-601-75
16	110	100.0	2488	22	US-10-741-600-244
17	110	100.0	3522	24	US-10-450-763-22266
18	110	100.0	6510	20	US-10-741-601-72
19	110	100.0	6510	22	US-10-741-600-241
20	110	100.0	7361	19	US-10-236-392-3
21	110	100.0	7677	22	US-10-956-157-4995
22	110	100.0	7795	16	US-10-084-817-2
23	110	100.0	7823	20	US-10-741-601-77
24	110	100.0	7823	22	US-10-741-600-245
25	110	100.0	7848	20	US-10-741-601-78
26	110	100.0	7848	22	US-10-741-600-246
27	110	100.0	7867	14	US-10-098-841-6
28	110	100.0	7935	20	US-10-741-601-74
29	110	100.0	7935	22	US-10-741-600-240
30	110	100.0	7959	20	US-10-741-601-81
31	110	100.0	7959	22	US-10-741-600-249
32	110	100.0	8013	20	US-10-741-601-71
33	110	100.0	8013	22	US-10-741-600-242
34	110	100.0	8027	18	US-10-447-161-8
35	110	100.0	8027	20	US-10-734-564-27
36	110	100.0	8027	22	US-10-852-335A-53
37	110	100.0	8027	24	US-10-287-436A-81
38	110	100.0	8044	17	US-10-240-965-121
39	110	100.0	8044	24	US-10-765-700-135
40	110	100.0	8062	14	US-10-098-841-5
41	110	100.0	8137	14	US-10-098-841-8
42	110	100.0	8155	20	US-10-741-601-79
43	110	100.0	8155	22	US-10-741-600-247
44	110	100.0	8226	20	US-10-741-601-69
45	110	100.0	8226	22	US-10-741-600-237

ALIGNMENTS

RESULT 1

US-09-765-231A-1
; Sequence 1, Application US/09765231A
; Patent No. US20020119452A1
; GENERAL INFORMATION:
; APPLICANT: Searle/Monsanto
; APPLICANT: Philippard, Deborah
; APPLICANT: Vaseanthakum, Geetha
; APPLICANT: Dotson, Stanton
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Osteoarthritis tissue-derived nucleic acids, polypeptides,
; TITLE OF INVENTION: vectors, and cells
; FILE REFERENCE: SO-3221 PR
; CURRENT APPLICATION NUMBER: US/09/765.231A

```
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 82
; SEQ ID NO 1
; LENGTH: 310
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-231A-1

Alignment Scores:
Pred. No.: 1.58e-10 Length: 310
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-8 (1-20) x US-09-765-231A-1 (1-310)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 19 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCAC 78

RESULT 2
US-10-344-634-3
; Sequence 3, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Kitajima, Tetsuya
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human Fibronectin
; OTHER INFORMATION: Collagen-binding Domain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)..(693)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (109)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (206)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (270)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (374)
; FEATURE:
; NAME/KEY: conflict
; LOCATION: (681)
US-10-344-634-3

Alignment Scores:
Pred. No.: 4.11e-10 Length: 708
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-09-581-651D-8 (1-20) x US-10-344-634-3 (1-708)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 352 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCAC 411

RESULT 3
US-10-198-846-8918
; Sequence 8918, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8918
; LENGTH: 731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 3, 5, 6, 7, 12, 716
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-8918

Alignment Scores:
Pred. No.: 4.26e-10 Length: 731
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-581-651D-8 (1-20) x US-10-198-846-8918 (1-731)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 481 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCAC 540

RESULT 4
US-10-344-634-15
; Sequence 15, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 20.1307 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-8

Perfect score: 110

Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool h/US09581651/runat_07112005_092224_28801/app_query.fasta_1.1834
-DB=Issued_Patents_NA -QMI=fastcap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 @cgn 1 1 187 @runat_07112005_092224_28801 -NCPUP=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	8044	4	US-09-566-921-135
2	106	96.4	7679	4	Sequence 135, App
3	106	96.4	7680	4	Sequence 38, Appl
4	106	96.4	7680	5	Sequence 1289, Ap
5	106	96.4	7705	1	PCT-US95-09819-6
6	106	96.4	7705	1	Sequence 16, Appl
7	106	96.4	7705	6	Sequence 16, Appl
8	106	96.4	7705	6	Patent No. 5455158
9	106	96.4	7803	2	Patent No. 5455158
10	106	96.4	7803	5	Sequence 1, Appl
11	105	95.5	986	1	PCT-US93-12887-1
12	105	95.5	986	1	Sequence 8, Appl

13	55	50.0	2109	4	US-09-799-451-345	Sequence 345, App
14	55	50.0	2334	1	US-08-457-304A-33	Sequence 33, Appl
15	55	50.0	2334	1	US-08-456-701A-33	Sequence 33, Appl
16	55	50.0	2334	3	US-08-684-932A-33	Sequence 33, Appl
17	55	50.0	2334	4	US-09-023-655-996	Sequence 996, App
18	55	50.0	2334	4	US-09-949-016-704	Sequence 704, App
19	55	50.0	2335	4	US-09-799-451-346	Sequence 346, App
20	55	50.0	2335	4	US-09-949-016-4758	Sequence 4758, Ap
c 21	53	48.2	4810	3	US-09-596-824-5	Sequence 5, Appl
c 22	53	48.2	4810	4	US-09-885-329-5	Sequence 1, Appl
c 23	52	47.3	50937	3	US-09-428-517-1	Sequence 1, Appl
24	51	46.4	11665	4	US-09-949-016-12446	Sequence 12446, A
25	51	46.4	11665	4	US-09-949-016-16500	Sequence 16500, A
c 26	48	43.6	10445	4	US-09-949-016-12311	Sequence 12311, A
c 27	48	43.6	10445	4	US-09-949-016-12927	Sequence 12927, A
c 28	48	43.6	58844	4	US-09-949-016-13769	Sequence 13769, A
c 29	47	42.7	1164	4	US-09-902-540-350	Sequence 350, App
c 30	47	42.7	1434	4	US-09-248-796A-1775	Sequence 1775, Ap
c 31	47	42.7	2157	4	US-09-657-960-2	Sequence 2, Appl
c 32	47	42.7	10302	4	US-09-657-960-1	Sequence 1, Appl
c 33	47	42.7	25356	4	US-09-976-594-750	Sequence 750, App
c 34	47	42.7	106746	4	US-09-326-402C-1	Sequence 1, Appl
c 35	47	42.7	106746	4	US-09-326-402C-12	Sequence 12, Appl
c 36	47	42.7	162465	4	US-09-949-016-14264	Sequence 14264, A
c 37	47	42.7	4403765	3	US-09-103-840A-2	Sequence 2, Appl
c 38	47	42.7	4411529	3	US-09-103-840A-1	Sequence 1, Appl
c 39	46	41.8	204	4	US-09-252-991A-9996	Sequence 9996, Ap
c 40	46	41.8	422	4	US-09-270-767-6921	Sequence 6921, Ap
c 41	46	41.8	422	4	US-09-270-767-22203	Sequence 22203, A
c 42	46	41.8	601	4	US-09-949-016-32909	Sequence 32909, A
c 43	46	41.8	601	4	US-09-949-016-67287	Sequence 67287, A
c 44	46	41.8	601	4	US-09-949-016-67288	Sequence 67288, A
c 45	46	41.8	1560	4	US-09-724-678D-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Alignment Scores:
Pred. No.: 8.65e-09 Length: 8044
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-8 (1-20) x US-09-566-921-135 (1-8044)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyehis 20

Db 1472 ACAGACCACACTGTTTGGTTTCAGACTCGAGGAGAAATCCAAATGGTCTTGCCAC 1531

RESULT 2
US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:
Pred. No.: 4,148-08 Length: 7679
Score: 106.00 Matches: 19
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.00% Mismatches: 0
Query Match: 96.36% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-8 (1-20) x US-09-220-132-38 (1-7679)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCyHis 20
Db 1124 ACAGACCACACTGTTTGGTTCAGACTCAAGGAGGAAATTCCAATGGTGCTTGTGCCAC 1183

RESULT 3
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; EXPRESSION
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g31396
US-09-023-655-1289

Alignment Scores:
Pred. No.: 4,148-08 Length: 7680
Score: 106.00 Matches: 19
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 95.00% Mismatches: 0
Query Match: 96.36% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-8 (1-20) x US-09-023-655-1289 (1-7680)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCyHis 20
Db 1124 ACAGACCACACTGTTTGGTTCAGACTCAAGGAGGAAATTCCAATGGTGCTTGTGCCAC 1183

RESULT 4
PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US95-09819-6
Alignment Scores:
Pred. No.: 4,148-08 Length: 7680

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 65.5686 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-8
Perfect score: 110
Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO.spool_h/US09581651/runat_07112005_092222_28772/app_query.fasta_1.1834
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 @CGN 1 1 1052 @runat_07112005_092222_28772 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Datbase : N Geneseq 16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	110	100.0	310	AAH23071	Aah23071 Osteoarth
2	110	100.0	706	ADSL19041	Adsl19041 Human fib
3	110	100.0	708	ABL50261	Abi50261 Human fib
4	110	100.0	731	ACN87768	Acn87768 Breast ca
5	110	100.0	741	AAL20536	Aal20536 Human bre

6	110	100.0	879	6	ABL50269	Abi50269 Collagen-
7	110	100.0	1051	13	ADSL19042	Adsl19042 Human fib
8	110	100.0	1053	3	AAA64263	Aaa64263 cDNA enco
9	110	100.0	1053	6	ABL41020	Abi41020 Modified
10	110	100.0	1182	6	ABL50268	Abi50268 Collagen-
11	110	100.0	1189	11	ACN90406	Acn90406 Breast ca
12	110	100.0	1224	3	AAA64271	Aaa64271 Fibronect
13	110	100.0	1407	6	ABL41027	Abi41027 Human fib
14	110	100.0	1443	6	ABL41028	Abi41028 Human fib
15	110	100.0	1527	3	AAA64270	Aaa64270 Fibronect
16	110	100.0	1929	13	ADSL17488	Adsl17488 Nucleotid
17	110	100.0	1929	13	ADR97657	Adr97657 Human fib
18	110	100.0	2127	10	ADD18477	Add18477 Human pro
19	110	100.0	2147	2	AAX81299	Aax81299 Human mlg
20	110	100.0	2358	13	ADR67201	Adr67201 Human bla
21	110	100.0	2443	13	ADQ38575	Adq38575 Human SNP
22	110	100.0	2488	13	ADQ38581	Adq38581 Human SNP
23	110	100.0	2856	13	ADSL19043	Adsl19043 Chimeric
24	110	100.0	3522	5	AAS86462	Aas86462 DNA encod
25	110	100.0	3540	12	ADP75954	Adp75954 Human min
26	110	100.0	3669	12	ADP29134	Adp29134 Human sec
27	110	100.0	4080	12	ADP75958	Adp75958 Human leu
28	110	100.0	6510	13	ADQ38578	Adq38578 Human SNP
29	110	100.0	6816	10	ABX74443	Abx74443 Human CDN
30	110	100.0	7242	13	ADR66637	Adr66637 Human pro
31	110	100.0	7242	13	ADR65953	Adr65953 Human pro
32	110	100.0	7361	9	ACD06170	Acc06170 Human cDN
33	110	100.0	7550	8	ACC00412	Acc00412 Human cel
34	110	100.0	7677	13	ACN37820	Acn37820 Tumour-as
35	110	100.0	7795	10	ADJ56196	Adj56196 Zebrafish
36	110	100.0	7823	13	ADQ38582	Adq38582 Human SNP
37	110	100.0	7848	13	ADQ38583	Adq38583 Human SNP
38	110	100.0	7867	4	AAI57803	Aai57803 Human pol
39	110	100.0	7935	13	ADQ38577	Adq38577 Human SNP
40	110	100.0	7959	13	ADQ38586	Adq38586 Human SNP
41	110	100.0	8013	13	ADQ38579	Adq38579 Human SNP
42	110	100.0	8027	11	ADP64998	Adp64998 Human fib
43	110	100.0	8027	12	ADG89565	Adg89565 Human fib
44	110	100.0	8027	12	ADQ29601	Adq29601 Human col
45	110	100.0	8027	13	ADR67200	Adr67200 Human bla

ALIGNMENTS

RESULT 1	
AAH23071	
ID	AAH23071 standard; DNA; 310 BP.
XX	
AC	AAH23071;
DT	17-SEP-2001 (first entry)
XX	
DE	Osteoarthritis tissue-derived nucleic acid sequence #1.
XX	
KW	Osteoarthritis; infectious disorder; Crohn's disease; sepsis; human;
KW	wound healing; osteopathic; anti-arthritis; anti-inflammatory; vulnary;
KW	antibacterial; antiallergic; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200153531-A2.
XX	
PD	26-JUL-2001.
XX	
PF	18-JAN-2001; 2001WO-US0000016.
XX	
PR	18-JAN-2000; 2000US-0176523P.
XX	
PA	(PHAA) PHARMACIA CORP.
XX	
PI	Phippard D, Vasanthakamur G, Dotson S, Ma X;
XX	
DR	WPI; 2001-451914/48.

XX Substantially purified protein, polypeptide or their fragments, used to
PT identify a biologically active compound or composition and treat
PT mammalian osteoarthritis.
XX Claim 1; Page 90; 144pp; English.
XX Sequences AAH23071-23152 represent nucleic acid sequences derived from
CC osteoarthritis tissues. The sequences are useful as probes and for the
CC diagnosis or prognosis of mammalian osteoarthritis. The polynucleotides
CC and polypeptides of the invention are useful for generating diagnostic
CC reagents, as targets for small molecule drug development, generation of
CC therapeutic, and cloning genes. Specific antibodies are used to generate
CC enzyme linked immunosorbent assays for detection of osteoarthritis. The
CC invented molecules can be used to treat osteoarthritis or to analyse the
CC disease-modifying activity of osteoarthritis drugs. Other disorders
CC treatable using the nucleic acid sequences include atopic, inflammatory
CC and infectious disorders e.g. Crohn's disease and sepsis, and wound
CC healing
XX Sequence 310 BP; 87 A; 75 C; 77 G; 71 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.68e-09 Length: 310
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-581-651D-8 (1-20) x AAH23071 (1-310)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCyHis 20
Db 19 ACAGACCACACTGTTTGGTTTCAGACTCGAGGAGGAAATTCCAATGGTGGCTTGTGCCAC 78
RESULT 2
ADSI19041
ID ADSI19041 standard; DNA; 706 BP.
XX
AC ADSI19041;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human fibronectin collagen binding domain DNA SeqID 2.
XX
KW human; ds; hepatocyte growth factor; HGF; collagen binding domain;
KW fibronectin; artificial blood vessel; stent; tissue regeneration;
KW vulnerable.
XX
OS Homo sapiens.
XX
PN JP2004269423-A.
XX
PD 30-SEP-2004.
XX
PF 07-MAR-2003; 2003JP-00062169.
XX
PR 07-MAR-2003; 2003JP-00062169.
XX
PA (TERU) TERUMO CORP.
XX
XX WPI; 2004-682709/67.
XX
PT Hepatocyte growth factor HGF derivative for inducing angiogenesis,
PT comprises fusion protein containing HGF polypeptide and polypeptide other
PT than HGF, and polypeptide having collagen binding property connected
PT directly or through intron.
XX
PS Example 1; SEQ ID NO 2; 51pp; Japanese.
XX
CC This invention relates to a novel hepatocyte growth factor (HGF)
CC derivative. Specifically, it refers to a fusion protein containing an HGF

CC polypeptide and the collagen binding domain of fibronectin, where the
CC collagen binding property and the HGF activity are maintained. The
CC present invention describes this HGF derivative as a live organ
CC transplant material that can be an artificial blood vessel or a stent in
CC which the collagen material is bound on its surface in a sheet-like
CC molding. Accordingly, it is useful for inducing and enhancing
CC angiogenesis (compared to natural type HGF) in damaged tissue and thus
CC can enhance tissue regeneration. Furthermore, it exhibits vulnerary
CC activity with increased stability and collagen binding properties. This
CC polynucleotide sequence is the DNA encoding the human fibronectin
CC collagen binding domain from Ala260 to Ala484, given in an
CC exemplification of the invention.
XX
SQ Sequence 706 BP; 187 A; 181 C; 187 G; 151 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.6e-09 Length: 706
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-581-651D-8 (1-20) x ADSI19041 (1-706)
Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCyHis 20
Db 349 ACAGACCACACTGTTTGGTTTCAGACTCGAGGAGGAAATTCCAATGGTGGCTTGTGCCAC 408
RESULT 3
ABL50261
ID ABL50261 standard; cDNA; 708 BP.
XX
AC ABL50261;
XX
DT 13-JUN-2002 (first entry)
XX
DE Human fibronectin collagen-binding domain encoding cDNA SEQ ID NO:3.
XX
KW Human; fibronectin; collagen-binding domain; collagen-binding hybrid;
KW basic fibroblast growth factor; enterokinase recognition site; collagen;
KW epidermal growth factor; drug delivery system; tissue regeneration; Gene;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 16..702
FT /*tag= a
FT /product= "human fibronectin collagen-binding domain"
XX
XX WO200214505-A1.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-JP007036.
XX
XX 15-AUG-2000; 2000JP-00246341.
XX
XX (TERU) TERUMO CORP.
XX
XX Ishikawa T, Kitajima T;
XX
XX WPI; 2002-257605/30.
XX
XX P-PSDB; ABB06747.
XX
XX Collagen-binding hybrid polypeptide, useful in drug delivery system for
XX functional polypeptides, formulating into complex with collagen to give
XX functionally modified collagen matrix as biomaterial for tissue
XX regeneration.
XX
XX Example 1; Page 59-61; 86pp; Japanese.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 535.477 Seconds
(without alignment)
1809.796 Million cell updates/sec

Title: US-09-581-651D-8
Perfect score: 110
Sequence: 1 TDHTVLVQTRGNSNGALCH 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=GenEmbl -PMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651@cgn 1 1 9336 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	310	6	AX202071 Sequence
2	110	100.0	615	9	AF312399 Homo sapi
3	110	100.0	708	6	BD142777 Collagen-
4	110	100.0	741	6	CQ427970 Sequence
c					

5	110	100.0	879	6	BD142785
6	110	100.0	1053	6	BD133458
7	110	100.0	1053	6	BD133469
8	110	100.0	1053	6	E63263
9	110	100.0	1182	6	BD142784
10	110	100.0	1224	6	E63271
11	110	100.0	1407	6	BD133465
12	110	100.0	1428	6	BD133476
13	110	100.0	1443	6	BD133466
14	110	100.0	1527	6	E63270
15	110	100.0	1560	6	BD133477
16	110	100.0	1929	6	CQ871810
17	110	100.0	1929	6	CQ871828
18	110	100.0	2147	6	BD137021
19	110	100.0	2147	6	AX003229
20	110	100.0	2147	9	HA276395
21	110	100.0	2192	9	HA535086
22	110	100.0	2358	6	CQ875358
23	110	100.0	2409	6	CQ731571
24	110	100.0	4321	9	HSM806214
25	110	100.0	6816	6	CQ715726
26	110	100.0	7299	9	HSM807162
27	110	100.0	7323	11	BV178337
28	110	100.0	7502	9	HSM804082
29	110	100.0	7544	9	HSM803509
30	110	100.0	7753	9	AB191261
31	110	100.0	7868	9	HSM806653
32	110	100.0	7951	9	HSM806901
33	110	100.0	7951	9	HSM806902
34	110	100.0	8027	6	CQ833991
35	110	100.0	8027	6	CQ875357
36	110	100.0	8030	9	HSM806170
37	110	100.0	8035	9	HSM806171
38	110	100.0	8042	9	HSM806805
39	110	100.0	8044	6	AR454662
40	110	100.0	8044	6	AX281712
41	110	100.0	8320	9	HSM806257
42	110	100.0	8411	9	HSM806992
43	110	100.0	8421	9	HSM806903
44	106	96.4	1139	9	HSU42404
45	106	96.4	7679	6	AR274901

ALIGNMENTS

RESULT 1	AX202071	Sequence 1 from Patent WO0153531.	310 bp	DNA	linear	PAT 30-AUG-2001
LOCUS	AX202071	Sequence 1 from Patent WO0153531.				
DEFINITION	AX202071	Sequence 1 from Patent WO0153531.				
ACCESSION	AX202071	Sequence 1 from Patent WO0153531.				
VERSION	AX202071.1	GI:15391856				
KEYWORDS	AX202071.1	GI:15391856				
SOURCE	AX202071.1	GI:15391856				
ORGANISM	AX202071.1	GI:15391856				
REFERENCE	AX202071.1	GI:15391856				
AUTHORS	AX202071.1	GI:15391856				
TITLE	AX202071.1	GI:15391856				
JOURNAL	AX202071.1	GI:15391856				
FEATURES	AX202071.1	GI:15391856				
source	AX202071.1	GI:15391856				
ORIGIN	AX202071.1	GI:15391856				
Alignment Scores:	AX202071.1	GI:15391856				
Pred. No.:	AX202071.1	GI:15391856				
Score:	AX202071.1	GI:15391856				
Percent Similarity:	AX202071.1	GI:15391856				

Alignment Scores: 1.43e-08 Length: 310
Pred. No.: 110.00 Matches: 20
Score: 100.00% Conservative: 0

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-8 (1-20) x AX202071 (1-310)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 19 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCCAATGGTGCCTTGTGCCAC 78

RESULT 2
AF312399 AF312399 615 bp mRNA linear PRI 07-NOV-2000
LOCUS Homo sapiens fibronectin mRNA, partial cds.
DEFINITION AF312399
ACCESSION AF312399
VERSION AF312399.1 GI:11119230
KEYWORDS
ORGANISM Homo sapiens (human)
SOURCE
REFERENCE 1 (bases 1 to 615)
AUTHORS Kornblith,A.R., Umezawa,K., Vibe-Pedersen,K. and Baralle,F.E.
TITLE Primary structure of human fibronectin: differential splicing may
generate at least 10 polypeptides from a single gene
EMBO J. 4 (7), 1755-1759 (1985)
JOURNAL
MEDLINE 85284965
PUBMED 2992939
REFERENCE 2 (bases 1 to 615)
AUTHORS Steffensen,B. and Martin,P.A.
TITLE Fibronectin collagen binding domain (modules I6, I11, I12, I7)
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 615)
AUTHORS Steffensen,B. and Martin,P.A.
TITLE Direct Submision
JOURNAL Submitted (10-OCT-2000) Periodontics, University of Texas Health
Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio,
TX 78229-3900, USA

FEATURES
source
1..615
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2q34"
/cell_lines="gingival fibroblasts IMR90"
/notes="identical to sequences found in gingival fibroblast
cell-lines WI38, MRC5, and Hs578T"
<1..615

CDS
/notes="spans modules I6, I11, I12, and I7; collagen
binding domain"
/codon_start=1
/product="fibronectin"
/protein_id="ANG30571.1"
/db_xref="GI:11119231"
/translation="GHCVTDSGVYVSGVQMLKTQGNKQMLCTCLGNVGSQETAVTQ
TYGNSGPECVLPFTYNGRTFYSCCTEGRODHLWCSTTSNYEQDKYSECTDHTVL
VQTRGNGNGLCHPFLYNNHNYTDCTSEGRDNNKWCCTQNYDQKFGFCPMAA
HBEICTTNEGVYRIGDQWDQHDHGMHMRCTCVGNRGREWTCTA"
52
conflict
/citation=[1]
/replace="t"
149
conflict
/citation=[1]
/replace="t"
213
conflict
/citation=[1]
/replace="g"
317
conflict
/citation=[1]
/replace="a"

ORIGIN
1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
352 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCCAATGGTGCCTTGTGCCAC 411

Alignment Scores:
Pred. No.: 2.66e-08 Length: 615
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-8 (1-20) x AF312399 (1-615)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 289 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCCAATGGTGCCTTGTGCCAC 348

RESULT 3
BD142777
LOCUS BD142777 708 bp DNA linear PAT 18-SEP-2002
DEFINITION Collagen-binding hybrid polypeptide.
ACCESSION BD142777
VERSION BD142777.1 GI:23237722
KEYWORDS WO 0214505-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 708)
AUTHORS Ishikawa,T. and Kitajima,T.
TITLE Collagen-binding hybrid polypeptide
JOURNAL Patent: WO 0214505-A 3 21-FEB-2002;
TERUMO CORP.TETSUYA ISHIKAWA,TAKASHI KITAJIMA
OS Artificial Sequence
PN WO 0214505-A/3
PD 21-FEB-2002
PF 15-AUG-2001 WO 2001JP007036
PR 15-AUG-2000 JP 00P 246341
PI TETSUYA ISHIKAWA,TAKASHI KITAJIMA
PC C12N15/09,C07K14/78,C07K19/00,C12N1/21,C12N5/10,A61K38/00 CC
Description of Artificial Sequence: Human
Fibronectin Collagen
binding
CC Domain Location/Qualifiers
FH Key (16)..(693) Location/Qualifiers
FT CDS (109)
FT conflict (206)
FT conflict (270)
FT conflict (374)
FT conflict (681)
FEATURES
source
1..708
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Alignment Scores:
Pred. No.: 3.02e-08 Length: 708
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-8 (1-20) x BD142777 (1-708)

Qy 1 ThrAspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCyHis 20
Db 352 ACAGACCACACTGTTTGGTTCAGACTCGAGGAGGAATTCCAATGGTGCCTTGTGCCAC 411

RESULT 4
CQ427970/c
LOCUS CQ427970 741 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 13004 from Patent WO0151628.
ACCESSION CQ427970
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 7.06536 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95.5	74.0	749	2 Q7Z391	Q7Z391 homo sapien
2	95.5	74.0	1103	2 Q6MZ4	Q6MZ4 homo sapien
3	95.5	74.0	2193	2 Q6MZ7	Q6MZ7 homo sapien
4	95.5	74.0	2240	2 Q6MDP8	Q6MDP8 homo sapien
5	95.5	74.0	2267	2 Q6MDP9	Q6MDP9 homo sapien
6	95.5	74.0	2296	2 Q6N0A6	Q6N0A6 homo sapien
7	95.5	74.0	2357	2 Q6MDT4	Q6MDT4 homo sapien
8	95.5	74.0	2386	1 F1NC HUMAN	P02751 homo sapien
9	95.5	74.0	2444	2 Q6N025	Q6N025 homo sapien
10	95.5	74.0	2477	1 F1NC MOUSE	P11276 mus musculus
11	95.5	74.0	2477	2 Q6MZU5	Q6MZU5 homo sapien
12	92.5	71.7	2265	1 F1NC BOVIN	P07589 bos taurus
13	88.5	68.6	2477	1 F1NC RAT	P04937 rattus norv
14	71.5	55.4	712	1 MM09 BOVIN	P52176 bos taurus
15	67.5	52.3	324	2 Q9N282	Q9N282 bos taurus
16	65.5	50.8	704	1 MM09 CANFA	O18733 canis fami
17	64.5	50.0	208	2 Q9XSD1	Q9XSD1 canis fami
18	64.5	50.0	704	2 Q71U09	Q71U09 canis fami
19	64.5	50.0	2319	2 Q9SMJ4	Q9SMJ4 myotis luci
20	63.5	49.2	661	2 Q9GLE5	Q9GLE5 bos taurus
21	62.5	48.4	670	2 Q6DF16	Q6DF16 xenopus tro
22	62.5	48.4	586	2 Q9DE15	Q9DE15 gallus gall
23	62.5	48.4	2328	2 Q9SMI9	Q9SMI9 erinaceus e
24	61.5	47.7	671	2 Q6PF33	Q6PF33 xenopus lae
25	61.5	47.7	730	2 Q9W7L6	Q9W7L6 xenopus lae
26	61.5	47.7	730	1 MM09 MOUSE	P41245 mus musculus
27	61.5	47.7	730	2 Q80X18	Q80X18 mus musculus
28	61.5	47.7	2421	2 Q9SMJ1	Q9SMJ1 lemur catta
29	61.5	47.7	2481	1 F1NC_XENLA	Q91740 xenopus lae
30	61.5	47.7	2481	2 Q6GQAS	Q6GQAS xenopus lae
31	60.5	46.9	1463	1 PA2R_BOVIN	P49259 bos taurus

32 60.5 46.9 2420 2 Q95M18
33 59.5 46.1 248 2 Q9BGL4
34 59.5 46.1 1326 2 Q13019
35 59.5 46.1 1458 1 PA2R_RABIT
36 59.5 46.1 1465 2 Q13018
37 58.5 45.3 708 1 MM09_RAT
38 58.5 45.3 2437 2 Q95MJ3
39 58.5 45.3 2473 2 Q95LC7
40 58.5 45.3 2487 2 Q9N1T0
41 58.5 45.3 2499 1 MPRI_BOVIN
42 57.5 44.6 2358 2 Q95MJ2
43 57.5 44.6 2425 2 Q95MJ0
44 56.5 43.8 385 2 Q9TUL8
45 56.5 43.8 632 2 Q9N1P6

ALIGNMENTS

RESULT 1
Q7Z391
ID Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Human colon endothel primary cell culture;
RA Bloecher H., Bocher M., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -
DR HSSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDP97F3ED2F0DE CRC64;

Query Match 74.0%; Score 95.5; DB 2; Length 749;

Best Local Similarity 55.3%; Pred. No. 3.1e-06;

Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
DB 451 PCVLPFTYNGRTFTYCTTEGQDGHLMCSTTSNYEQDQ 488

RESULT 2

Q6MZ4
ID Q6MZ4 PRELIMINARY; PRT; 1103 AA.
AC Q6MZ4;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DE 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219 (Fragment)).

```
GN Name=DKFZp686K139; Synonyms=DKFZp686P219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansoerg W., Krieger S., Regliert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE4200.1; -.
DR EMBL; BX640802; CAE45885.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FECACAF634AD56 CRC64;

Query Match 74.0%; Score 95.5; DB 2; Length 1103;
Best Local Similarity 55.3%; Pred. No. 4.9e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPTFYNDRT-----DSTTSNYEQDQ 23
DB 448 PCVLPTFYNGRTFYCTTEGRQDGHLCSTTSNYEQDQ 485

RESULT 3
Q6MZM7 PRELIMINARY; PRT; 2193 AA.
AC Q6MZM7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686O12165 (Fragment).
GN Name=DKFZp686O12165;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG The German Human cDNA Consortium;
RA Lauber J., Bahr A., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G.,
RA Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640999; CAE46002.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.

GN Name=DKFZp686K139; Synonyms=DKFZp686P219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Human cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
```

```
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 7.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 7.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 17.
DR PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01253; FIBRONECTIN_1; 7.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 17.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 2193 AA; 240641 MW; F876E93106540EF3 CRC64;

Query Match 74.0%; Score 95.5; DB 2; Length 2193;
Best Local Similarity 55.3%; Pred. No. 1.1e-05;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPTFYNDRT-----DSTTSNYEQDQ 23
DB 106 PCVLPTFYNGRTFYCTTEGRQDGHLCSTTSNYEQDQ 143

RESULT 4
Q68DP8 PRELIMINARY; PRT; 2240 AA.
AC Q68DP8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686H0342.
GN Name=DKFZp686H0342;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749317; CAH18172.1; -.
DR InterPro; IPR002086; Aldehyd dehydrog.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 10.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.74379 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLFFYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	2386	1 FNHU	fibronectin precu
2	92.5	71.7	2265	1 FNBO	fibronectin - bovi
3	88.5	68.6	2477	2 S14428	fibronectin precu
4	71.5	55.4	712	1 I46031	gelatinase B (EC 3
5	61.5	47.7	730	1 I52580	gelatinase B (EC 3
6	61.5	47.7	730	2 JCI456	gelatinase B (EC 3
7	61.5	47.7	2481	2 A43908	fibronectin - Afri
8	60.5	46.9	1463	2 A53210	phospholipase A2 r
9	59.5	46.1	1326	2 B56395	secretory phosphol
10	59.5	46.1	1458	1 A49707	phospholipase A2 r
11	59.5	46.1	1465	2 A56395	secretory phosphol
12	58.5	45.3	708	2 J43664	gelatinase B (EC 3
13	58.5	45.3	708	2 S62907	gelatinase B (EC 3
14	58.5	45.3	2499	1 A30788	mannose 6-phosphat
15	56.5	43.8	660	1 A28153	gelatinase A (EC 3
16	55.5	43.0	662	2 A42496	gelatinase A (EC 3
17	55.5	43.0	662	2 S34780	gelatinase A (EC 3
18	55.5	43.0	2491	1 A28372	insulin-like growt
19	54.5	42.2	603	2 S28941	coagulation factor
20	52.5	40.7	662	2 S70365	gelatinase A (EC 3
21	52.5	40.7	707	1 A53795	gelatinase B (EC 3
22	52.5	40.7	2482	2 I48922	cation-independent
23	52.5	40.7	2483	1 A49617	insulin-like growt
24	50.5	39.1	1487	2 S48719	insulin-like growt
25	50	38.8	456	2 S45137	fushi tarazu segme
26	49.5	38.4	707	1 A34458	gelatinase B (EC 3
27	49	38.0	326	2 A47523	cartilage homeopro
28	49	38.0	2242	2 A57541	pyrimidine synthe
29	48.5	37.6	593	2 S45281	coagulation factor

ALIGNMENTS

RESULT 1

FNHU

fibronectin precursor [validated] - human
N/Alternate names: fibronectin splice form ED-A

C/Species: Homo sapiens (man)

C/Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C/Accession: A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R/Dean: D.C.; Bowlus, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A/Reference number: A26460; MUID:87175578; PMID:3031656

A/Accession: A26460

A/Molecule type: DNA

A/Residues: 1-49 <DEA>

A/Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:G182686; PIDN:AAA5337

R/Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A/Title: Evolution of the fibronectin gene.

A/Reference number: A26284; MUID:86111901; PMID:3003095

A/Accession: A26284

A/Molecule type: DNA

A/Residues: 1447-1540 <OLD>

A/Cross-references: GB:M12549; NID:G182688

A/Note: the authors translated the codon TTC for residue 1494 as Glu

R/Paoletti, G.; Henschliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 15, 3545-3557, 1988

A/Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A/Reference number: S00848; MUID:86233940; PMID:3375063

A/Accession: S03917

A/Molecule type: DNA

A/Residues: 1594-1767, 'V', 1769-1783 <PAO>

A/Cross-references: EMBL:X07718; NID:G31402

A/Note: the authors translated the codon AAC for residue 1631 as Asp

R/Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A/Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:

A/Reference number: A24854; MUID:87030929; PMID:3770201

A/Accession: A24854

A/Molecule type: DNA

A/Residues: 1992-2147 <VIB>

A/Cross-references: GB:X04530; NID:G31436

R/Gutman, A.; Yamada, K.M.; Kornblitt, A.

FEBS Lett. 207, 145-148, 1986

A/Title: Human fibronectin is synthesized as a pre-polypeptide.

A/Reference number: A24476; MUID:87030890; PMID:3770189

A/Accession: A24476

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-14, 'Q', 16-38 <GUT>

R/Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.

EMBO J. 4, 1753-1759, 1985

A/Title: Primary structure of human fibronectin: differential splicing may generate at

transcription regu
hepatocyte growth
probable dimethyl
hemocytin - silkwo
coagulation factor
mannose receptor,
hypothetical prote
C-8 sterol isomera
seminal fluid prot
flagellar M-ring p
F83508
hypothetical prote
AD3-like protein
hypothetical prote
hypothetical prote
zonadhesin - mouse

30 48 37.2 461 2 C86679
31 48 37.2 655 1 A46688
32 48 37.2 812 2 A90680
33 48 37.2 3133 2 S52093
34 47.5 36.8 615 1 KFHU12
35 47.5 36.8 1479 2 T42710
36 47 36.4 166 2 T18691
37 47 36.4 222 2 JH0488
38 46.5 36.0 115 2 A29156
39 46.5 36.0 598 2 F83508
40 46 35.7 124 2 A84463
41 46 35.7 325 2 T33711
42 46 35.7 551 2 T40053
43 46 35.7 764 2 H71607
44 46 35.7 884 2 T40690
45 46 35.7 5376 2 T42215

A:Reference number: A91008; MUID:85284965; PMID:2992939
A:Accession: A91008
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 32-1344, 1346-2080;2112-2386 <KOR>
A:Cross-references: GB:X02761
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A:Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A:Reference number: A93529; MUID:84272258; PMID:6462219
A:Accession: A93529
A:Molecule type: mRNA
A:Residues: 973-2080;2112-2386 <K02>
A:Cross-references: GB:X00739
R:Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A:Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
A:Reference number: A21011; MUID:83290929; PMID:6688418
A:Accession: A21011
A:Molecule type: mRNA
A:Residues: 1434-1537 <OL2>
A:Cross-references: GB:K00055; NID:g182680; PIDN:AAAS2459.1; PID:g182683
R:Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A:Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A:Reference number: A90495; MUID:85280409; PMID:2992573
A:Accession: A90495
A:Molecule type: mRNA
A:Residues: 1594-2386 <BER>
A:Cross-references: GB:M10905; NID:g182696; PIDN:AAAS2462.1; PID:g182697
R:Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A:Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A:Reference number: A22245; MUID:85231203; PMID:2989004
A:Accession: A22245
A:Molecule type: mRNA
A:Residues: 1948-2067 <UM5>
A:Cross-references: GB:M27589; NID:g182705; PIDN:AAAS2465.1; PID:g182706
A:Accession: B22245
A:Molecule type: mRNA
A:Residues: 1975-1991;2017-2039 <UM2>
A:Cross-references: GB:M27590
R:Sekiguchi, K.; Klos, A.M.; Kurechi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A:Title: Human liver fibronectin complementary DNAs: identification of two different me
A:Reference number: I52394; MUID:87026578; PMID:3021206
A:Accession: I65273
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
A:Cross-references: GB:M14060; NID:g182701; PIDN:AAAS2464.1; PID:g182704
R:Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A:Reference number: A21165; MUID:83221567; PMID:6304659
A:Accession: A21165
A:Molecule type: mRNA
A:Residues: 2291-2386 <K03>
A:Cross-references: GB:K00799; NID:g182681; PIDN:AAAS2460.1; PID:g182684
R:Garcia-Pardo, A.; Pearlestein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A:Title: Primary structure of human plasma fibronectin.
A:Reference number: A92398; MUID:84032463; PMID:6630202
A:Accession: A92398
A:Molecule type: protein
A:Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>
R:Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A:Title: Further characterization of the binding of fibronectin to gelatin reveals the p
A:Reference number: S34791; MUID:93312001; PMID:8323285
A:Accession: S34791
A:Molecule type: protein
A:Residues: 291-300;551-560 <GAR2>

R:Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A:Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A:Reference number: A60904; MUID:87019725; PMID:3532418
A:Accession: A60904
A:Molecule type: protein
A:Residues: 293-301 <GR1>
R:Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A:Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A:Reference number: A23901; MUID:86008277; PMID:3900070
A:Accession: A23901
A:Molecule type: protein
A:Residues: 616-677, 'Q', 679-703, 'P', 'R' <CAL>
R:Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A:Title: The cell attachment domain of fibronectin. Determination of the primary structu
A:Reference number: A92386; MUID:82265604; PMID:7050098
A:Accession: A92386
A:Molecule type: protein
A:Residues: 1441-1548 <PIE>
A:Note: residues 1524-1527 are responsible for the cell-binding activity
R:Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A:Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A:Reference number: A32517; MUID:87241275; PMID:3593230
A:Accession: A32517
A:Molecule type: protein
A:Residues: 1599-1630, 'T', 1722-2058 <GAR3>
R:Tresselt, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A:Title: Human plasma fibronectin. Demonstration of structural differences between the A
A:Reference number: S14357; MUID:91190085; PMID:2012601
A:Accession: S14357
A:Molecule type: protein
A:Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>
R:Garcia-Pardo, A.; Pearlestein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A:Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A:Reference number: A23891; MUID:85261459; PMID:4019516
A:Accession: A23891
A:Molecule type: protein
A:Residues: 2071-2080;2112-2356 <GAR4>
A:Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C:Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
ation, and transformation.
C:Genetics:
A:Gene: GDB:FNI
A:Cross-references: GDB:119135; OMIM:135600
A:Map position: 2q34-2q34
A:Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
A:Superfamily: fibronectin; fibronectin type I repeat homology <1F1>
C:Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-31/Domain: propeptide #status predicted <PRO>
F:32-2386/Product: fibronectin #status experimental
F:52-272/Domain: fibrin and heparin binding <FHB>
F:52-87/Domain: fibronectin type I repeat homology <1F1>
F:97-135/Domain: fibronectin type I repeat homology <1F2>
F:141-179/Domain: fibronectin type I repeat homology <1F3>
F:186-225/Domain: fibronectin type I repeat homology <1F4>
F:231-270/Domain: fibronectin type I repeat homology <1F5>
F:308-342/Domain: collagen binding <CBR>
F:360-401/Domain: fibronectin type II repeat homology <2F1>
F:420-461/Domain: fibronectin type II repeat homology <2F2>
F:470-508/Domain: fibronectin type I repeat homology <1F7>
F:518-555/Domain: fibronectin type I repeat homology <1F8>
F:561-599/Domain: fibronectin type I repeat homology <1F9>
F:609-692/Domain: fibronectin type III repeat homology <3FA>
F:616-706/Domain: heparin binding <HPB>
F:719-801/Domain: fibronectin type III repeat homology <3FB>
F:810-891/Domain: fibronectin type III repeat homology <3FC>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 7.78693 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	113	87.6	642	16	US-10-741-601-354
2	113	87.6	642	17	US-10-741-600-1066
3	95.5	74.0	228	15	US-10-344-634-4
4	95.5	74.0	285	15	US-10-344-634-16
5	95.5	74.0	343	9	US-09-934-706-1
6	95.5	74.0	386	15	US-10-344-634-14
7	95.5	74.0	400	9	US-09-934-706-5
8	95.5	74.0	463	15	US-10-144-194A-52
9	95.5	74.0	463	16	US-10-491-566-52
10	95.5	74.0	501	9	US-09-934-706-4
11	95.5	74.0	657	16	US-10-741-601-359

12	95.5	74.0	657	17	US-10-741-600-1072	Sequence 1072, Ap
13	95.5	74.0	984	16	US-10-741-601-356	Sequence 356, App
14	95.5	74.0	984	17	US-10-741-600-1069	Sequence 1069, App
15	95.5	74.0	1173	18	US-10-450-763-52634	Sequence 52634, A
16	95.5	74.0	2220	15	US-10-236-392-4	Sequence 4, Appli
17	95.5	74.0	2296	16	US-10-741-601-363	Sequence 363, App
18	95.5	74.0	2326	17	US-10-741-600-1075	Sequence 1075, Ap
19	95.5	74.0	2326	14	US-10-279-733-8	Sequence 8, Appli
20	95.5	74.0	2320	15	US-10-236-392-2	Sequence 2, Appli
21	95.5	74.0	2328	14	US-10-171-311-64	Sequence 64, Appl
22	95.5	74.0	2328	15	US-10-236-031B-70	Sequence 70, Appl
23	95.5	74.0	2328	15	US-10-374-979-98	Sequence 98, Appl
24	95.5	74.0	2328	15	US-10-182-336A-98	Sequence 98, Appl
25	95.5	74.0	2328	16	US-10-477-238A-677	Sequence 677, App
26	95.5	74.0	2328	16	US-10-680-287A-677	Sequence 677, App
27	95.5	74.0	2328	17	US-10-477-173-677	Sequence 677, App
28	95.5	74.0	2355	15	US-10-144-194A-104	Sequence 104, App
29	95.5	74.0	2355	15	US-10-360-101-235	Sequence 235, App
30	95.5	74.0	2355	15	US-10-447-161-3	Sequence 3, Appli
31	95.5	74.0	2355	16	US-10-734-564-94	Sequence 94, Appl
32	95.5	74.0	2355	16	US-10-741-601-357	Sequence 357, App
33	95.5	74.0	2355	16	US-10-741-601-366	Sequence 366, App
34	95.5	74.0	2355	16	US-10-491-566-104	Sequence 104, App
35	95.5	74.0	2355	17	US-10-741-600-1067	Sequence 1067, Ap
36	95.5	74.0	2355	17	US-10-741-600-1078	Sequence 1078, Ap
37	95.5	74.0	2355	17	US-10-852-335A-147	Sequence 147, App
38	95.5	74.0	2355	18	US-10-287-436A-436	Sequence 436, App
39	95.5	74.0	2355	18	US-10-287-436A-1137	Sequence 1137, Ap
40	95.5	74.0	2355	20	US-11-040-130-28	Sequence 28, Appl
41	95.5	74.0	2386	10	US-09-961-403-1	Sequence 1, Appli
42	95.5	74.0	2386	16	US-10-741-601-360	Sequence 360, App
43	95.5	74.0	2386	16	US-10-788-792-206	Sequence 206, App
44	95.5	74.0	2386	16	US-10-618-281-32	Sequence 32, Appl
45	95.5	74.0	2386	17	US-10-741-600-1071	Sequence 1071, Ap

ALIGNMENTS

RESULT 1
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 87.6%; Score 113; DB 16; Length 642;
Best Local Similarity 91.3%; Pred. No. 4.1e-08;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVLPFTYNDRTDSTTSNYEQDQ 23
DB 359 PCVLPFTYNGHTCTTSNYEQDQ 381

RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match      87.6%; Score 113; DB 17; Length 642;
Best Local Similarity 91.3%; Pred. No. 4.1e-08;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PCVLPFTYNDRTDSTTSNYEQDQ 23
Db 359 PCVLPFTYNGRTCTTSNYEQDQ 381

RESULT 3
US-10-344-634-4
; Sequence 4, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Human Fibronectin
; OTHER INFORMATION: Collagen-binding Domain
US-10-344-634-4

Query Match      74.0%; Score 95.5; DB 15; Length 228;
Best Local Similarity 55.3%; Pred. No. 5.3e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
Db 70 PCVLPFTYNGRTFYSCCTTEGRQDGHLMCSTTSNYEQDQ 107

RESULT 4
US-10-344-634-16
; Sequence 16, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 16
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Collagen-binding
US-10-344-634-16

Query Match      74.0%; Score 95.5; DB 15; Length 285;
Best Local Similarity 55.3%; Pred. No. 6.8e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
Db 70 PCVLPFTYNGRTFYSCCTTEGRQDGHLMCSTTSNYEQDQ 107

RESULT 5
US-09-934-706-1
; Sequence 1, Application US/09934706
; Patent No. US20020102709A1
; GENERAL INFORMATION:
; APPLICANT: Terumo Corporation
; TITLE OF INVENTION: Functional Hybrid Polypeptide with Collagen-binding
; TITLE OF INVENTION: Activity
; FILE REFERENCE: 19990120
; CURRENT APPLICATION NUMBER: US/09/934,706
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified Human
; OTHER INFORMATION: Fibronectin Collagen-Binding Domain
; NAME/KEY: INIT_MET
; LOCATION: (1)
; NAME/KEY: DOMAIN
; LOCATION: (2)...(341)
; OTHER INFORMATION: /note="human fibronectin collagen-binding domain"
; NAME/KEY: CONFLICT
; LOCATION: (69)
; NAME/KEY: CONFLICT
; LOCATION: (125)
US-09-934-706-1

Query Match      74.0%; Score 95.5; DB 9; Length 343;
Best Local Similarity 55.3%; Pred. No. 8.4e-06;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 PCVLPFTYNDRT-----DSTTSNYEQDQ 23
Db 70 PCVLPFTYNGRTFYSCCTTEGRQDGHLMCSTTSNYEQDQ 107

RESULT 6
US-10-344-634-14
; Sequence 14, Application US/10344634
; Publication No. US20040053368A1
; GENERAL INFORMATION:
; APPLICANT: Ishikawa, Tetsuya
; APPLICANT: Kitajima, Takashi
; TITLE OF INVENTION: Collagen-Binding Hybrid Polypeptide
; FILE REFERENCE: 029650-132
; CURRENT APPLICATION NUMBER: US/10/344,634
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: PCT/JP01/07036
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: JP 2000-246341
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.43529 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLPTFYNDRTSTTSNYEQDQ 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCJTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	60	2	US-08-982-597A-18
2	95.5	74.0	60	3	US-09-136-218-26
3	95.5	74.0	2231	1	US-08-153-799-16
4	95.5	74.0	2324	1	US-08-283-857-1
5	95.5	74.0	2324	5	PCT-US95-09819-1
6	95.5	74.0	2327	6	5455158-1
7	95.5	74.0	2327	6	5455158-1
8	95.5	74.0	2386	2	US-09-016-366A-12
9	95.5	74.0	2386	4	US-09-961-403-1
10	95.5	74.0	2446	2	US-08-551-356-2
11	95.5	74.0	2446	5	PCT-US93-12687-2
12	60.5	46.9	1463	1	US-08-220-603A-11
13	59.5	46.1	1326	4	US-09-949-016-6806
14	59.5	46.1	1339	4	US-09-949-016-10448
15	59.5	46.1	1478	4	US-09-949-016-8315
16	56.5	43.8	429	4	US-09-194-468A-45
17	56.5	43.8	631	3	US-08-448-489-17
18	56.5	43.8	631	4	US-09-689-730-17
19	56.5	43.8	660	3	US-08-704-711A-18
20	56.5	43.8	660	3	US-09-521-220-18
21	56.5	43.8	660	3	US-09-391-104-19
22	56.5	43.8	660	4	US-09-917-254-89
23	56.5	43.8	660	4	US-09-949-016-6512
24	56.5	43.8	660	4	US-09-949-016-7937
25	55.5	43.0	2491	4	US-09-207-363-1
26	54	41.9	24	2	US-08-982-597A-23
27	54	41.9	24	3	US-09-136-218-23

28	52	40.3	16	2	US-08-982-597A-26	Sequence 26, Appl
29	52	40.3	16	3	US-09-136-218-26	Sequence 26, Appl
30	50.5	39.1	65	2	US-08-982-597A-19	Sequence 19, Appl
31	50.5	39.1	65	3	US-09-136-218-19	Sequence 19, Appl
32	50.5	39.1	1487	3	US-08-840-062-7	Sequence 7, Appl
33	49.5	38.4	707	3	US-08-704-711A-19	Sequence 19, Appl
34	49.5	38.4	707	3	US-09-521-220-19	Sequence 19, Appl
35	49.5	38.4	707	3	US-09-391-104-20	Sequence 20, Appl
36	49.5	38.4	707	4	US-09-949-016-6575	Sequence 6575, Ap
37	49.5	38.4	708	3	US-08-448-489-16	Sequence 16, Appl
38	49.5	38.4	708	4	US-09-689-730-16	Sequence 16, Appl
39	49.5	38.4	713	4	US-09-949-016-10629	Sequence 10629, A
40	49	38.0	373	4	US-08-248-796A-18953	Sequence 18953, A
41	48	37.2	410	1	US-08-091-519-13	Sequence 13, Appl
42	48	37.2	410	1	US-08-442-043A-13	Sequence 13, Appl
43	48	37.2	410	3	US-09-173-151A-27	Sequence 27, Appl
44	48	37.2	410	4	US-09-461-908-13	Sequence 13, Appl
45	48	37.2	410	4	US-08-441-893A-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-982-597A-18
; Sequence 18, Application US/08982597A
; Patent No. 5932693
; GENERAL INFORMATION:
; APPLICANT: Santoro, Samuel A.
; APPLICANT: Staats, William D.
; TITLE OF INVENTION: Antithrombotic Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer
; STREET: 800 No. 5932693th Lindbergh Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,597A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,542
; FILING DATE: 10-DEC-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-3002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-694-3117
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-982-597A-18

Query Match 74.0%; Score 95.5; DB 2; Length 60;
Best Local Similarity 55.3%; Pred. No. 1.7e-07;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 PCVLPTFYNDRT-----DSTTSNYEQDQ 23
15 PCVLPTFYNGRTFYSCCTTEGQDGLMCGTTSNYEQDQ 52
DB

RESULT 2
US-09-136-218-18
; Sequence 18, Application US/09136218
; Patent No. 6083914
; GENERAL INFORMATION:
; APPLICANT: Santoro, Samuel A.
; APPLICANT: Staatz, William D.
; TITLE OF INVENTION: Antithrombotic Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer
; STREET: 800 No. 6083914th Lindbergh Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: USA
; ZIP: 63167

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09136,218
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/982,597
FILING DATE:
APPLICATION NUMBER: 60/032,542
FILING DATE: 10-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-3002
TELEPHONE: 314-694-3117
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-136-218-18

Query Match 74.0%; Score 95.5; DB 3; Length 60;
Best Local Similarity 55.3%; Pred. No. 1.7e-07;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 PCVLPTNYNDRT-----DSTTSNYEQDQ 23
Db 15 PCVLPTNYGRFYSCCTTEGRQDGHLMWCSTTSNYEQDQ 52

RESULT 3
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..2231
OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match 74.0%; Score 95.5; DB 1; Length 2231;
Best Local Similarity 55.3%; Pred. No. 1e-05;
Matches 21; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 PCVLPTNYNDRT-----DSTTSNYEQDQ 23
Db 328 PCVLPTNYGRFYSCCTTEGRQDGHLMWCSTTSNYEQDQ 365

RESULT 4
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5752742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 8.5085 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129

Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

- 1: geneseq1980s.*
- 2: geneseq1990s.*
- 3: geneseq2000s.*
- 4: geneseq2001s.*
- 5: geneseq2002s.*
- 6: geneseq2003as.*
- 7: geneseq2003bs.*
- 8: geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	129	100.0	23	2	AAY28912
2	129	100.0	24	2	AAY28904
3	129	100.0	660	2	AAY28901
4	122	94.6	642	8	ADR67316
5	122	94.6	642	8	ADS17489
6	122	94.6	642	8	ADR97658
7	113	87.6	642	8	ADQ39403
8	95.5	74.0	39	2	AAY28908
9	95.5	74.0	60	2	AAW62371
10	95.5	74.0	228	5	ABB06747
11	95.5	74.0	285	5	ABB06751
12	95.5	74.0	343	3	ABB08505
13	95.5	74.0	343	5	ABB07961
14	95.5	74.0	386	5	ABB06750
15	95.5	74.0	400	3	ABB08509
16	95.5	74.0	461	5	ABB07964
17	95.5	74.0	463	6	ABR58303
18	95.5	74.0	473	5	ABB07965
19	95.5	74.0	501	3	ABB08508
20	95.5	74.0	657	8	ADQ39409
21	95.5	74.0	720	2	AAY28914
22	95.5	74.0	951	8	ADS19044
23	95.5	74.0	984	8	ADQ39406
24	95.5	74.0	1173	4	ABG22275
25	95.5	74.0	1179	8	ADP75952

26	95.5	74.0	1223	8	ADP30365	Adp30365 Human sec
27	95.5	74.0	1359	8	ADP75957	Adp75957 Human leu
28	95.5	74.0	2182	8	ADR66462	Adr66462 Human pro
29	95.5	74.0	2182	8	ADR66120	Adr66120 Human pro
30	95.5	74.0	2220	6	ABO01289	AbO01289 Human pro
31	95.5	74.0	2220	8	ADN95950	Adn95950 Human NOV
32	95.5	74.0	2265	4	AAM38647	Aam38647 Human pol
33	95.5	74.0	2266	6	ABR40124	AbR40124 Human cel
34	95.5	74.0	2296	8	ADQ39412	AdQ39412 Human myo
35	95.5	74.0	2320	6	AAE37107	Aae37107 Human fib
36	95.5	74.0	2320	6	ABR42588	AbR42588 Human fib
37	95.5	74.0	2320	6	ABO01288	AbO01288 Human pro
38	95.5	74.0	2320	8	ADN95948	Adn95948 Human NOV
39	95.5	74.0	2320	8	ADR90519	Adr90519 Human ful
40	95.5	74.0	2324	2	AAR92778	Aar92778 Human fib
41	95.5	74.0	2324	5	AAU74674	Aau74674 Human fib
42	95.5	74.0	2324	5	AAE23651	Aae23651 Human pro
43	95.5	74.0	2327	1	AAP70373	Aap70373 Human fib
44	95.5	74.0	2327	2	AAR15468	Aar15468 Human fib
45	95.5	74.0	2328	4	AAG68182	Aag68182 Fibronect

ALIGNMENTS

RESULT 1

AAY28912
ID AAY28912 standard; peptide; 23 AA.

AC AAY28912;

DT 21-SBP-1999 (first entry)

DE MSF 1-alpha peptide epitope.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.

OS Synthetic.

OS Homo sapiens.

FN WO9931233-A1.

PD 24-JUN-1999.

PF 15-DEC-1998; 98WO-GB003766.

PR 16-DEC-1997; 97GB-00026539.

PA (UYDU-) UNIV DUNDEE.

PI Schor SL, Schor AM;

XX WPI; 1999-430039/36.

Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

Claim 29; Page 61; 86pp; English.

The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28912-913 CC represent peptide epitopes of MSF against which monoclonal antibodies can be raised

SQ Sequence 23 AA;

Query Match 100.0%; Score 129; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 1.3e-12;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 478.34 Seconds

(without alignment)
1830.243 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTSTTSNYEQDQ 23

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651.QCGN 1 1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_hc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gest1:
9: gb_gest2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	380	CN701692	CN701692 E0452H02-
2	95.5	74.0	385	1 AL589418	AL589418 DKF2p451P
3	95.5	74.0	392	2 BE079849	BE079849 RC6-BT062
4	95.5	74.0	400	2 BE122858	BE122858 14.08 Hum
5	95.5	74.0	411	5 BX642740	BX642740 DKF2p781G
6	95.5	74.0	431	2 BE815722	BE815722 PM3-BN017
7	95.5	74.0	431	6 CD550770	CD550770 B0318C06-
8	95.5	74.0	458	1 AJ661698	AJ661698 AJ661698
9	95.5	74.0	471	4 BG007031	BG007031 RC1-GN023

10	95.5	74.0	471	5	BX495442	BX495442 DKF2p7791
11	95.5	74.0	475	4	BG007376	BG007376 IL5-GN024
12	95.5	74.0	478	5	BX495600	BX495600 DKF2p779E
13	95.5	74.0	484	4	BF986449	BF986449 QV4-GN012
14	95.5	74.0	505	7	CF166960	CF166960 B0778B11-
15	95.5	74.0	507	4	BF985411	BF985411 QV4-GN014
16	95.5	74.0	510	2	AW368063	AW368063 CM0-HT018
17	95.5	74.0	513	5	BX642768	BX642768 DKF2p781H
18	95.5	74.0	521	1	AJ668111	AJ668111 AJ668111
19	95.5	74.0	528	7	CF793207	CF793207 886238 MA
20	95.5	74.0	538	1	AJ685866	AJ685866 AJ685866
21	95.5	74.0	552	4	BG945246	BG945246 PM0-AN008
22	95.5	74.0	558	7	CN419467	CN419467 170005325
23	95.5	74.0	564	2	BF229278	BF229278 PM0-AN008
24	95.5	74.0	598	5	BQ340570	BQ340570 PM0-NN025
25	95.5	74.0	605	7	CN481619	CN481619 hw10a12.Y
26	95.5	74.0	618	7	CN482377	CN482377 hw19f03.Y
27	95.5	74.0	620	7	CN481412	CN481412 hw07d04.Y
28	95.5	74.0	622	7	CN419625	CN419625 170005326
29	95.5	74.0	626	7	CF907680	CF907680 A0513D09-
30	95.5	74.0	635	5	BQ312550	BQ312550 PM0-BN017
31	95.5	74.0	645	4	BG945249	BG945249 PM0-AN008
32	95.5	74.0	645	7	CN419499	CN419499 170004551
33	95.5	74.0	646	2	AW321646	AW321646 uo36f10.Y
34	95.5	74.0	655	7	CN419546	CN419546 170005328
35	95.5	74.0	657	7	CN419503	CN419503 170004706
36	95.5	74.0	702	7	CN419472	CN419472 170005326
37	95.5	74.0	705	1	AU140899	AU140899 UI-H-E21-
38	95.5	74.0	725	5	BQ574857	BQ574857 UI-H-E21-
39	95.5	74.0	748	1	AU140948	AU140948 AU140948
40	95.5	74.0	748	1	AU140991	AU140991 AU140991
41	95.5	74.0	758	1	AU140556	AU140556 AU140556
42	95.5	74.0	759	1	AU140735	AU140735 AU140735
43	95.5	74.0	780	7	CN162843	CN162843 952371 MA
44	95.5	74.0	781	7	CN161442	CN161442 950735 MA
45	95.5	74.0	786	4	BI149026	BI149026 602311306

ALIGNMENTS

RESULT 1

CN701692

LOCUS

DEFINITION

musculus cDNA clone NIA: E0452H02 IMAGE:30873589 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

CN701692 E0452H02-5 NIA Mouse E11.5 whole embryo cDNA library (Long) Mus
musculus cDNA clone NIA: E0452H02 IMAGE:30873589 5', mRNA sequence.

CN701692 380 bp mRNA linear EST 18-MAY-2004

CN701692.1 GI:47470441

EST.
Mus musculus (house mouse)

1 (bases 1 to 380),
Sharov,A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y.,
VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Baesey,U.C.,
Wang,X., Carter,M.G., Hamatani,T., Alba,K., Akutsu,H., Sharova,L.,
Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S.,
Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L.,
Schlessinger,D., Keller,J., Klotz,E., Kelsae,G., Umezawa,A.,
Verscoi,A.L., Rossant,J., Kunath,T., Hogan,B.L., Curci,A.,
D'Urso,M., Kelsae,J., Hide,W. and Ko,M.S.

Transcriptome analysis of mouse stem cells and early embryos
PLOS Biol. 1 (3), 410-419 (2003)

Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igun.grc.nia.nih.gov

Plate: E0452 row: H column: 02
Seq primer: M13 Reverse
High quality sequence stop: 380
POLYA-No.

Location/Qualifiers

FEATURES

```

source
1..380
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="niaEST:E0452H02-5"
/db_xref="taxon:10090"
/clone="NIA:E0452H02 IMAGE:30873589"
/tissue_type="whole embryo including extraembryonic
tissues at 11.5-days postcoitum"
/dev_stage="E11.5"
/lab_host="DH10B"
/clone_lib="NIA Mouse E11.5 whole embryo cDNA library
(long)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgeun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were extracted from a pool of 3 embryos at 11.5-days
postcoitum. Double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen]:
5'-pGACTAGTCTAGATCGGAGCGCGCCCTTTT-3' from
2ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.3Kb. The library was
constructed by Yulan Piao."

ORIGIN
Alignment Scores:
Pred. No.: 7,91e-06 Length: 380
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
DB: 7 Gaps: 1

US-09-581-651D-7 (1-23) x CN701692 (1-380)

Qy 1 ProCysValLeuProPheThrTyRAsnAspArgThr----- 12
Db 219 CCCTGTGCTCTCCCTGACCAACGGTAGGACCTTCTATTCCTGCACCCAGGAGG 278
Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 279 CGGCAAGACGGACATCTGTGGTGTAGCACCACTTCATTCAGAACAGACCAG 332

RESULT 2
AL589418 385 bp mRNA linear EST 04-SEP-2003
LOCUS DKEP2451P1415.r1 451 (synonym: hlcc1) spinal cord Homo sapiens cDNA
DEFINITION clone DKEP2451P1415 5', mRNA sequence.
ACCESSION AL589418
VERSION AL589418.1 GI:13243190
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 385)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,W., Weil,B. and
Wiemann,S.
TITLE EST (Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Weil,B. and
Wiemann,S.)

Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKF2p451P1415) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..385
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p451P1415"
/tissue_type="human spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="451 (synonym: hlcc1) spinal cord"
/notes="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"

ORIGIN
Alignment Scores:
Pred. No.: 8.03e-06 Length: 385
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
DB: 1 Gaps: 1

US-09-581-651D-7 (1-23) x AL589418 (1-385)

Qy 1 ProCysValLeuProPheThrTyRAsnAspArgThr----- 12
Db 133 CCATGTGCTTACCATTCACCTACCAATGGCAGGAGCTTCTACTCTGCACCCAGGAGG 192
Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 193 CGACAGGACGGACATCTTTGTGTGCAGCACCACTTCGAATTATGACGAGCCAG 246

RESULT 3
BE079849/c 392 bp mRNA linear EST 12-JUN-2000
LOCUS RC6-BT0627-220300-012-F02 BT0627 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE079849
VERSION BE079849.1 GI:8470133
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Goldman,G.H., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
Brunstein,A., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
O'Hare,M.J.
Simpson,A.J.
Shogun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL 20202663
MEDLINE 10737800
PUBMED 10737800
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 118.097 Seconds
(without alignments)
1610.594 Million chr updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDSTTSNYEQDQ 23

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1824
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09581651@cgn1.1.1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	113	87.6	2127	17	US-10-210-120-49
2	113	87.6	2127	22	US-10-956-157-4288
3	113	87.6	2127	24	US-10-909-035-49
4	113	87.6	2443	20	US-10-741-601-70
5	113	87.6	2443	22	US-10-741-600-238
6	95.5	74.0	504	9	US-09-815-343-273
7	95.5	74.0	504	19	US-10-097-105-273
8	95.5	74.0	708	19	US-10-344-634-3
9	95.5	74.0	731	15	US-10-198-846-8918
10	95.5	74.0	879	19	US-10-344-634-15
11	95.5	74.0	1053	9	US-09-934-706-8
12	95.5	74.0	1182	19	US-10-344-634-13
13	95.5	74.0	1189	15	US-10-198-846-11556
14	95.5	74.0	1224	9	US-09-934-706-16
15	95.5	74.0	1527	9	US-09-934-706-15
16	95.5	74.0	2488	20	US-10-741-601-75
17	95.5	74.0	2488	22	US-10-741-600-244
18	95.5	74.0	3522	24	US-10-450-763-22266
19	95.5	74.0	4295	18	US-10-144-1948-51
20	95.5	74.0	4295	21	US-10-491-566-51
21	95.5	74.0	6510	20	US-10-741-601-72
22	95.5	74.0	6510	22	US-10-741-600-241
23	95.5	74.0	6988	19	US-10-236-392-1
24	95.5	74.0	7361	19	US-10-236-392-3
25	95.5	74.0	7677	22	US-10-956-157-4995
26	95.5	74.0	7679	22	US-10-831-704-38
27	95.5	74.0	7680	9	US-09-964-8244-574
28	95.5	74.0	7680	15	US-10-171-311-63
29	95.5	74.0	7680	18	US-10-236-031B-69
30	95.5	74.0	7680	19	US-10-374-979-75
31	95.5	74.0	7680	19	US-10-182-936A-75
32	95.5	74.0	7680	19	US-10-641-643-1289
33	95.5	74.0	7680	20	US-10-717-592-222
34	95.5	74.0	7680	20	US-10-788-777-79
35	95.5	74.0	7680	21	US-10-477-238A-654
36	95.5	74.0	7680	21	US-10-680-287A-654
37	95.5	74.0	7680	22	US-10-278-698-88
38	95.5	74.0	7680	22	US-10-478-698-603
39	95.5	74.0	7680	22	US-10-843-641A-5877
40	95.5	74.0	7680	22	US-10-477-173-654
41	95.5	74.0	7680	22	US-10-852-335A-52
42	95.5	74.0	7705	18	US-10-447-161-4
43	95.5	74.0	7795	16	US-10-084-817-2
44	95.5	74.0	7823	20	US-10-741-601-77
45	95.5	74.0	7823	22	US-10-741-600-245

ALIGNMENTS

RESULT 1
US-10-210-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UN-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02

```
/ PRIOR APPLICATION NUMBER: US 60/334,468
/ PRIOR FILING DATE: 2001-11-15
/ NUMBER OF SEQ ID NOS: 123
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 49
/ LENGTH: 2127
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-210-120-49

Alignment Scores:
Pred. No.: 5,23e-09 Length: 2127
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 17 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-210-120-49 (1-2127)

Qy 1 ProCysValLeuProPheThrTyraAsnAspArgThrAspSerThrThrSerAsnTyrGlu 20
Db 1131 CCATGTGCTTACCATTACCTACATGCGCAGGACGTGCAGCAACACTTCGAATTATGAG 1190

Qy 21 GlnAspGln 23
Db 1191 CAGGACCAG 1199

RESULT 2
US-10-956-157-4288
/ Sequence 4288, Application US/10956157
/ Publication No. US20050118625A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Mounts, William
/ TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
/ FILE REFERENCE: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
/ FILE REFERENCE: 031896-043000 (AM 101081)
/ CURRENT APPLICATION NUMBER: US/10/956,157
/ CURRENT FILING DATE: 2004-10-04
/ NUMBER OF SEQ ID NOS: 319805
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 4288
/ LENGTH: 2127
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-956-157-4288

Alignment Scores:
Pred. No.: 5,23e-09 Length: 2127
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 22 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-956-157-4288 (1-2127)

Qy 1 ProCysValLeuProPheThrTyraAsnAspArgThrAspSerThrThrSerAsnTyrGlu 20
Db 1131 CCATGTGCTTACCATTACCTACATGCGCAGGACGTGCAGCAACACTTCGAATTATGAG 1190

Qy 21 GlnAspGln 23
Db 1191 CAGGACCAG 1199

RESULT 3
US-10-909-035-49
/ Sequence 49, Application US/10909035
/ Publication No. US20050136493A1
/ GENERAL INFORMATION:
/ APPLICANT: Rubin, Mark A.
/ APPLICANT: Chinnaiyan, Arul M.

/ APPLICANT: Laxman, Bharathi
/ APPLICANT: Sreekumar, Arun
/ TITLE OF INVENTION: AMACR Cancer Markers
/ FILE REFERENCE: UM-09098
/ CURRENT APPLICATION NUMBER: US/10/909,035
/ CURRENT FILING DATE: 2004-07-30
/ NUMBER OF SEQ ID NOS: 123
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 49
/ LENGTH: 2127
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-909-035-49

Alignment Scores:
Pred. No.: 5,23e-09 Length: 2127
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 24 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-909-035-49 (1-2127)

Qy 1 ProCysValLeuProPheThrTyraAsnAspArgThrAspSerThrThrSerAsnTyrGlu 20
Db 1131 CCATGTGCTTACCATTACCTACATGCGCAGGACGTGCAGCAACACTTCGAATTATGAG 1190

Qy 21 GlnAspGln 23
Db 1191 CAGGACCAG 1199

RESULT 4
US-10-741-601-70
/ Sequence 70, Application US/10741601
/ Publication No. US20040168519A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001500
/ CURRENT APPLICATION NUMBER: US/10/741,601
/ CURRENT FILING DATE: 2003-12-22
/ NUMBER OF SEQ ID NOS: 26415
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 70
/ LENGTH: 2443
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-741-601-70

Alignment Scores:
Pred. No.: 6,21e-09 Length: 2443
Score: 113.00 Matches: 21
Percent Similarity: 91.30% Conservative: 0
Best Local Similarity: 91.30% Mismatches: 2
Query Match: 87.60% Indels: 0
DB: 20 Gaps: 0

US-09-581-651D-7 (1-23) x US-10-741-601-70 (1-2443)

Qy 1 ProCysValLeuProPheThrTyraAsnAspArgThrAspSerThrThrSerAsnTyrGlu 20
Db 1447 CCATGTGCTTACCATTACCTACATGCGCAGGACGTGCAGCAACACTTCGAATTATGAG 1506

Qy 21 GlnAspGln 23
Db 1507 CAGGACCAG 1515

RESULT 5
US-10-741-600-238
/ Sequence 238, Application US/10741600
/ Publication No. US20050026169A1
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 23.1503 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTDTSNVEQDQ 23

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092224_28801/app_query.fasta_1.1834
-DB=Issued Patents NA -QWTF=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCPU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651 -CGEN 1 1 187 @runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95.5	74.0	7679	4	US-09-220-132-38
2	95.5	74.0	7680	4	Sequence 38, Appl Sequence 1289, Ap
3	95.5	74.0	7680	5	PCT-US95-09819-6
4	95.5	74.0	7705	1	US-08-259-589-16
5	95.5	74.0	7705	2	US-08-826-885-16
6	95.5	74.0	7705	6	US455158-2
7	95.5	74.0	7705	6	Patent No. 5455158
8	95.5	74.0	7803	2	US-08-551-356-1
9	95.5	74.0	7803	5	PCT-US93-12687-1
10	95.5	74.0	8044	4	US-09-566-921-135
11	88.5	68.6	986	1	US-07-637-250A-8
12	88.5	68.6	986	1	US-08-145-061-8

13	60.5	46.9	4978	1	US-08-220-603A-1	Sequence 1, Appl
14	59.5	46.1	601	4	US-09-949-016-36296	Sequence 36296, A
15	59.5	46.1	601	4	US-09-949-016-86573	Sequence 86573, A
16	59.5	46.1	601	4	US-09-949-016-163266	Sequence 163266,
17	59.5	46.1	4621	4	US-09-949-016-4577	Sequence 4577, Ap
18	59.5	46.1	4627	4	US-09-949-016-935	Sequence 935, App
19	59.5	46.1	5627	4	US-09-949-016-2444	Sequence 2444, Ap
20	59.5	46.1	5633	4	US-09-023-655-1490	Sequence 1490, Ap
21	59.5	46.1	119930	4	US-09-949-016-12677	Sequence 12677, A
22	59.5	46.1	119931	4	US-09-949-016-16319	Sequence 16319, A
23	59.5	46.1	125536	4	US-09-949-016-14186	Sequence 14186, A
24	56.5	43.8	601	4	US-09-949-016-30321	Sequence 30321, A
25	56.5	43.8	601	4	US-09-949-016-70165	Sequence 70165, A
26	56.5	43.8	1983	4	US-09-949-016-2066	Sequence 2066, App
27	56.5	43.8	3069	4	US-09-949-016-641	Sequence 641, App
28	56.5	43.8	29954	4	US-09-949-016-13808	Sequence 13808, A
29	56.5	43.8	31040	4	US-09-949-016-12383	Sequence 12383, A
30	52	40.3	1383	4	US-08-248-796A-2647	Sequence 2647, Ap
31	50	38.8	601	4	US-09-949-016-42379	Sequence 42379, A
32	50	38.8	601	4	US-09-949-016-42380	Sequence 42380, A
33	50	38.8	601	4	US-09-949-016-96535	Sequence 96535, A
34	50	38.8	601	4	US-09-949-016-96801	Sequence 96801, A
35	50	38.8	601	4	US-09-949-016-97067	Sequence 97067, A
36	50	38.8	601	4	US-09-949-016-97333	Sequence 97333, A
37	50	38.8	601	4	US-09-949-016-97599	Sequence 97599, A
38	50	38.8	601	4	US-09-949-016-97865	Sequence 97865, A
39	50	38.8	601	4	US-09-949-016-98131	Sequence 98131, A
40	50	38.8	601	4	US-09-949-016-98397	Sequence 98397, A
41	50	38.8	601	4	US-09-949-016-98663	Sequence 98663, A
42	50	38.8	601	4	US-09-949-016-98929	Sequence 98929, A
43	50	38.8	601	4	US-09-949-016-99195	Sequence 99195, A
44	50	38.8	601	4	US-09-949-016-99461	Sequence 99461, A
45	50	38.8	601	4	US-09-949-016-99727	Sequence 99727, A

ALIGNMENTS

RESULT 1

US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:
Pred. No.: 9.42e-07
Score: 95.50
Percent Similarity: 55.26%
Best Local Similarity: 55.26%
Query Match: 74.03%
DB: 4
Matches: 21
Conservative: 0
Mismatch: 2
Indels: 15
Gaps: 1

US-09-581-651D-7 (1-23) x US-09-220-132-38 (1-7679)

Qy 1 ProCysValLeuProPheThrTyrAsnAspArgThr-----12

Db 995 CCATGTGTCTTACCATTCACCTACAAATGGCAGGACTTCTACTCTGCACCGAAGGG 1054

Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 1055 CGACGAGCGGACATCTTTGGTGCAGCACAACTTCGAATTATGAGCAGGACCAG 1108

RESULT 2
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: GENBANK
; US-09-023-655-1289

Alignment Scores:
Pred. No.: 9.42e-07 Length: 7680
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
DB: 4 Gaps: 1

US-09-581-651D-7 (1-23) x US-09-023-655-1289 (1-7680)

Qy 1 ProCysValLeuProPheThrTyrAsnAspArgThr----- 12
Db 995 CCATGTGCTTACCATTCACCTACATGCGCAGAGCTTCTACTCTCTGCACCGAAGG 1054

Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 1055 CGACGAGCGGACATCTTTGGTGCAGCACAACTTCGAATTATGAGCAGGACCAG 1108

RESULT 3

PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; PCT-US95-09819-6

Alignment Scores:
Pred. No.: 9.42e-07 Length: 7680
Score: 95.50 Matches: 21
Percent Similarity: 55.26% Conservative: 0
Best Local Similarity: 55.26% Mismatches: 2
Query Match: 74.03% Indels: 15
DB: 5 Gaps: 1

US-09-581-651D-7 (1-23) x PCT-US95-09819-6 (1-7680)

Qy 1 ProCysValLeuProPheThrTyrAsnAspArgThr----- 12
Db 995 CCATGTGCTTACCATTCACCTACATGCGCAGAGCTTCTACTCTCTGCACCGAAGG 1054

Qy 13 -----AspSerThrThrSerAsnTyrGluGlnAspGln 23
Db 1055 CGACGAGCGGACATCTTTGGTGCAGCACAACTTCGAATTATGAGCAGGACCAG 1108

RESULT 4
US-08-259-569-16
; Sequence 16, Application US/08259569
; Patent No. 5679320
; GENERAL INFORMATION:
; APPLICANT: Vogel, Tikva
; APPLICANT: Levanon, Avigdor
; APPLICANT: Werber, Moshe
; APPLICANT: Guy, Rachel
; APPLICANT: Panet, Amos
; APPLICANT: Hartman, Jacob
; APPLICANT: Shaked, Hadassa
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 75.4039 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-7

Perfect score: 129
Sequence: 1 PCVLPFTYNDRTSTTSNYEQDQ 23

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/Cgn2_1/USPTO.spool_h/US09581651/runat_07112005_092222_28772/app_query.fasta.1.1834
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1.1.1052 @runat_07112005_092222_28772 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	2147	2 AAX81299	Aax81299 Human mig
2	122	94.6	1929	13 ADS19041	Adel17488 Nucleotid
3	122	94.6	1929	13 ADR97657	Adr97657 Human fib
4	122	94.6	2358	13 ADR67201	Adr67201 Human bla
5	113	87.6	2127	10 ADD18477	Add18477 Human pro

6	113	87.6	2443	13	ADQ38575	Adq38575 Human SNP
7	95.5	74.0	504	4	AAS57597	Aas57597 cDNA #273
8	95.5	74.0	706	13	ADS19041	Adel19041 Human fib
9	95.5	74.0	708	6	ABL50261	Ab150261 Human fib
10	95.5	74.0	731	11	ACN87768	Acn87768 Breast ca
11	95.5	74.0	741	4	AAI20536	Aai20536 Human bre
12	95.5	74.0	879	6	ABL50269	Ab150269 Collagen-
13	95.5	74.0	1051	13	ADS19042	Adel19042 Human fib
14	95.5	74.0	1053	3	AAA64263	Aaa64263 cDNA enco
15	95.5	74.0	1053	6	ABL41020	Ab141020 Modified
16	95.5	74.0	1182	6	ABL50268	Ab150268 Collagen-
17	95.5	74.0	1189	11	ACN90406	Acn90406 Breast ca
18	95.5	74.0	1224	3	AAA64271	Aaa64271 Fibronect
19	95.5	74.0	1407	6	ABL41027	Ab141027 Human fib
20	95.5	74.0	1443	6	ABL41028	Ab141028 Human fib
21	95.5	74.0	1527	3	AAA64270	Aaa64270 Fibronect
22	95.5	74.0	2488	13	ADQ38581	Adq38581 Human SNP
23	95.5	74.0	2856	13	ADS19043	Adel19043 Chimeric
24	95.5	74.0	3522	5	AAS86462	Aas86462 DNA encod
25	95.5	74.0	3540	12	ADP75954	Adp75954 Human min
26	95.5	74.0	3669	12	ADP29134	Adp29134 Human sec
27	95.5	74.0	4080	12	ADP75958	Adp75958 Human leu
28	95.5	74.0	4295	8	ACC72037	Acc72037 BCU0770 g
29	95.5	74.0	4860	3	AAA35009	Aaa35009 Human ade
30	95.5	74.0	6510	13	ADQ38578	Adq38578 Human SNP
31	95.5	74.0	6816	10	ABX74443	Abx74443 Human CDN
32	95.5	74.0	6988	9	ACD06169	Acd06169 Human CDN
33	95.5	74.0	7020	13	ADR90518	Adr90518 Human CDN
34	95.5	74.0	7049	12	ADN95947	Adn95947 Human NOV
35	95.5	74.0	7242	13	ADR66637	Adr66637 Human pro
36	95.5	74.0	7242	13	ADR65953	Adr65953 Human pro
37	95.5	74.0	7361	9	ACD06170	Acd06170 Human CDN
38	95.5	74.0	7550	8	ACC00412	Acc00412 Human cel
39	95.5	74.0	7677	13	ACN37820	Acn37820 Tumour-as
40	95.5	74.0	7679	10	ADB31322	Adb31322 Testoster
41	95.5	74.0	7680	2	AAT17551	Aat17551 Human fib
42	95.5	74.0	7680	3	AAF21131	Aaf21131 Human low
43	95.5	74.0	7680	5	ABA82689	Aba82689 Fibronect
44	95.5	74.0	7680	6	ABL67540	Ab167540 Thyroid c
45	95.5	74.0	7680	6	ABT11082	Abt11082 Human bre

ALIGNMENTS

RESULT 1
AAX81299
ID AAX81299 standard; cDNA; 2147 BP.
XX
AC AAX81299;
XX
DT 21-SEP-1999 (first entry)
XX
DE Human migration stimulating factor (MSF) 1-alpha encoding cDNA.
XX
KW Migration stimulatory factor; MSF; cell migration; modulation; human;
XX wound healing; scarring; MSF1-alpha; ss.
XX Homo sapiens.
XX
FN WO9931233-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-GB003766.
XX
PR 16-DEC-1997; 97GB-00026539.
XX
PA (UYDU-) UNIV DUNDEE.
XX
PI Schor SL, Schor AM;
XX
DR WPI; 1999-430039/36.
DR P-PSDB; AAY28901.

```

XX  Proteins with cell migration stimulatory activity used in treating wound
PT  and preventing scarring.
XX
XX  Example 1; Fig 1; 86pp; English.
XX
XX  The invention provides a human migration stimulatory factor (MSF)
CC  protein. Host cells containing a replicable vector comprising the MSF
CC  encoding nucleic acid can be used for the recombinant production of the
CC  protein. The polypeptide can be used for modulating cell migration,
CC  healing a wound and for preventing scarring. The present sequence
CC  represents the nucleotide sequence encoding a human MSF1-alpha protein
XX
XX  Sequence 2147 BP; 575 A; 512 C; 575 G; 485 T; 0 U; 0 Other;
SQ

```

Alignment Scores:

Pred. No.:	6.34e-12	Length:	2147
Score:	129.00	Matches:	23
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-581-651D-7 (1-23) x AAX81299 (1-2147)

```

QY  1 ProCysValLeuProPheThrTyrAsnAspArgThrAspSerThrSerAsnTyrGlu 20
DB  1131 CCATGTGTCTTACCATTACCTACACGACGAGGACGACGACCACTTCGAATTATGAG 1190
QY  21 GlnAspGln 23
DB  1191 CAGGACCAG 1199

```

RESULT 2

ADSI17488

ID ADS17488 standard; DNA; 1929 BP.

XX ADS17488;

XX

DT 02-DEC-2004 (first entry)

XX

DE Nucleotide sequence of human fibronectin 1.

XX

KW cell state; time-lapse profile; protein-protein interaction;
 KW drug screening; cancer; infectious disease; allergy; hypertension;
 KW hyperlipaemia; diabetes; cardiac disease; cerebral infarction; dementia;
 KW obesity; arteriosclerosis; infertility; mental disease; nervous disease;
 KW cataract; progeria; hypersensitivity; ultraviolet radiation; human;
 KW fibronectin 1; actin acting substance; transfection array; gene; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..1929

FT /*tag= a

FT /product= "fibronectin 1"

XX

PN WO2004079007-A2.

XX

PD 16-SEP-2004.

XX

PF 03-MAR-2004; 2004WO-JP002694.

XX

PR 04-MAR-2003; 2003JP-00057870.

XX

PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX

PI Miyake M, Yoshikawa T, Uchimura E, Miyake J;

XX

DR WPI; 2004-662438/64.

XX

DR P-PSDB; ADS17489.

XX

PT Presenting a state of a cell, useful for diagnosing and treating a

```

PT  disease, e.g. cancer, infectious disease, allergy, diabetes, dementia,
PT  obesity, infertility, or cataract, comprises obtaining a time-lapse
XX  profile of the cell.
XX
XX  Disclosure; SEQ ID NO 1; 532pp; English.
XX
XX  The specification describes a method and system for accurately presenting
CC  a state of a cell. The method comprises obtaining a time-lapse profile of
CC  the cell by time-lapse monitoring of a gene state associated with at
CC  least one gene derived from the cell, and presenting the time-lapse
CC  profile. The gene comprises a transcription control sequence, and the
CC  gene state includes expression of the gene. The method and system are
CC  useful presenting a state of a cell. The method can allow the elucidation
CC  of key protein-protein interactions suitable for targeting by drug
CC  screening protocols. The method is useful for diagnosing or treating a
CC  disease, e.g. cancer, infectious disease due to viruses or bacteria,
CC  allergy, hypertension, hyperlipaemia, diabetes, cardiac disease, cerebral
CC  infarction, dementia, obesity, arteriosclerosis, infertility, mental and
CC  nervous diseases, cataract, progeria, or hypersensitivity to ultraviolet
CC  radiation. The present sequence encodes human fibronectin 1. Bovine
CC  fibronectin was used as a candidate for an actin acting substance. The
CC  actin acting substance was used with transfection reagents and amplified
CC  plasmid DNA in assays using transfection arrays, in the course of the
CC  invention.
XX
SQ  Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	8.95e-11	Length:	1929
Score:	122.00	Matches:	22
Percent Similarity:	95.65%	Conservative:	0
Best Local Similarity:	95.65%	Mismatches:	1
Query Match:	94.57%	Indels:	0
DB:	13	Gaps:	0

US-09-581-651D-7 (1-23) x ADS17488 (1-1929)

```

QY  1 ProCysValLeuProPheThrTyrAsnAspArgThrAspSerThrSerAsnTyrGlu 20
DB  1075 CCATGTGTCTTACCATTACCTACGAGGACGAGGACGACGACCACTTCGAATTATGAG 1134
QY  21 GlnAspGln 23
DB  1135 CAGGACCAG 1143

```

RESULT 3

ADR97657

ID ADR97657 standard; DNA; 1929 BP.

XX ADR97657;

XX

DT 02-DEC-2004 (first entry)

XX

DE Human fibronectin 1 DNA, an actin acting substance SeqID 1.

XX

KW human; gene; ds; transfection efficiency; actin acting substance;
 KW extracellular matrix; fibronectin 1; gene introduction reagent.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 1..1929

FT /*tag= a

FT /product= "Fibronectin protein"

XX

PN WO2004079332-A2.

XX

PD 16-SEP-2004.

XX

PF 03-MAR-2004; 2004WO-JP002696.

XX

PR 04-MAR-2003; 2003JP-00057869.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 615.799 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-7
Perfect score: 129
Sequence: 1 PCVLPFTYNDRTSTTSNYEQDQ 23

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=genEmbl -QFWT=fractap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651.qcgn 1.1 9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	2147	6	Bd137021 Polypepti
2	129	100.0	2147	6	AX003229 Sequence
3	129	100.0	2147	9	AJ276395 Homo sapi
4	122	94.6	1929	6	CQ871810 Sequence

ALIGNMENTS

```

RESULT 1
LOCUS      Bd137021
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION  Bd137021
VERSION     Bd137021.1 GI:23231966
KEYWORDS   JP 2002508179-A/1.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2147)
AUTHORS   Schor, S.L. and Schor, A.M.
TITLE      Polypeptides, polynucleotides and uses thereof
JOURNAL    Patent: JP 2002508179-A 1 19-MAR-2002;
            UNIVERSITY OF DUNDEE
COMMENT    OS Homo sapiens (human)
            PN JP 2002508179-A/1
            PD 19-MAR-2002
            PF 15-DEC-1998 JP 2000539133
            PR 16-DEC-1997 GB 9726539.1
            PI SETH LAWRENCE SCHOR, ANA MARIA SCHOR
            PC C12N15/09, A61K38/00, A61P17/02, C07K14/78, C07K16/18, C12N5/10, PC
            C12P21/08,
            PC C1201/68, G01N33/574, C12N15/00, A61K37/02, C12N5/00 CC
            Polypeptides, polynucleotides and uses thereof. FH Key
            Location/Qualifiers
            FT source
            1..2147

```

5	122	94.6	1929	6	CQ871828	Sequence
6	122	94.6	2358	6	CQ875358	Sequence
7	95.5	74.0	504	6	AX260622	Sequence
8	95.5	74.0	615	9	AF312399	Homo sapi
9	95.5	74.0	708	6	BD142777	Collagen-
10	95.5	74.0	741	6	CQ427970	Sequence
11	95.5	74.0	879	6	BD142785	Collagen-
12	95.5	74.0	1053	6	BD133458	Collagen-
13	95.5	74.0	1053	6	BD133469	Hybrid po
14	95.5	74.0	1053	6	E63263	Collagen-bi
15	95.5	74.0	1139	9	HSU42404	Human fibro
16	95.5	74.0	1182	6	BD142784	Collagen-
17	95.5	74.0	1224	6	E63271	Collagen-bi
18	95.5	74.0	1407	6	BD133465	Collagen-
19	95.5	74.0	1428	6	BD133476	Hybrid po
20	95.5	74.0	1443	6	BD133466	Collagen-
21	95.5	74.0	1527	6	E63270	Collagen-bi
22	95.5	74.0	1560	6	BD133477	Hybrid po
23	95.5	74.0	2192	9	HSAS35086	Sequence
24	95.5	74.0	2409	6	CQ731571	Homo sapi
25	95.5	74.0	3059	10	BC051082	Sequence
26	95.5	74.0	4321	9	BSM806214	Homo sapi
27	95.5	74.0	6816	6	CQ715726	Sequence
28	95.5	74.0	7299	9	BSM807162	Homo sapi
29	95.5	74.0	7323	11	BV178397	Sequence
30	95.5	74.0	7502	9	BSM804082	Homo sapi
31	95.5	74.0	7544	9	BSM803509	Homo sapi
32	95.5	74.0	7679	6	AR274901	Sequence
33	95.5	74.0	7680	6	AR380744	Sequence
34	95.5	74.0	7680	6	AX277596	Sequence
35	95.5	74.0	7680	6	AX335368	Sequence
36	95.5	74.0	7680	9	HSF1B1	Human mRNA
37	95.5	74.0	7705	6	A14133	Fibronectin
38	95.5	74.0	7705	6	AR034630	Sequence
39	95.5	74.0	7705	6	E01162	cDNA encodi
40	95.5	74.0	7705	6	I70110	Sequence 16
41	95.5	74.0	7705	6	AR364992	Sequence
42	95.5	74.0	7803	9	AB191261	Homo sapi
43	95.5	74.0	7803	6	AR051657	Sequence
44	95.5	74.0	7868	9	BSM806653	Homo sapi
45	95.5	74.0	7951	9	BSM806901	Homo sapi

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 6.45098 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	122	100.0	296	2	Q8C6J7	Q8C6J7 mus musculus
2	122	100.0	749	2	Q7Z391	Q7Z391 homo sapien
3	122	100.0	1103	2	Q6MZP4	Q6MZP4 homo sapien
4	122	100.0	2267	2	Q6MDP9	Q6MDP9 homo sapien
5	122	100.0	2296	2	Q6N0A6	Q6N0A6 homo sapien
6	122	100.0	2357	2	Q68DT4	Q68DT4 homo sapien
7	122	100.0	2386	1	FINC HUMAN	P02751 homo sapien
8	122	100.0	2444	2	Q6N0Z5	Q6N0Z5 homo sapien
9	122	100.0	2477	1	FINC MOUSE	P11276 mus musculus
10	122	100.0	2477	1	FINC RAT	P04937 rattus norv
11	122	100.0	2477	2	Q6MZU5	Q6MZU5 homo sapien
12	117	95.9	2265	1	FINC BOVIN	P07589 bos taurus
13	114	93.4	215	2	Q6DD34	Q6DD34 xenopus lae
14	114	93.4	2481	1	FINC XENLA	Q91740 xenopus lae
15	114	93.4	2481	2	Q6GQZ5	Q6GQZ5 xenopus lae
16	113	92.6	2240	2	Q68DP8	Q68DP8 homo sapien
17	84	68.9	922	2	Q33405	Q33405 brachydanio
18	84	68.9	2478	2	Q33406	Q33406 brachydanio
19	75	61.5	100	1	FINC RABIT	Q28749 oryctolagus
20	75	61.5	111	2	Q862C9	Q862C9 bos taurus
21	75	61.5	163	2	Q9H382	Q9H382 homo sapien
22	75	61.5	211	2	Q7L553	Q7L553 homo sapien
23	75	61.5	216	2	Q99KD0	Q99KD0 mus musculus
24	75	61.5	268	2	Q6PJS5	Q6PJS5 homo sapien
25	75	61.5	358	2	Q7T152	Q7T152 brachydanio
26	75	61.5	412	2	Q71U44	Q71U44 bos taurus
27	75	61.5	810	2	Q6R3F3	Q6R3F3 mus musculus
28	75	61.5	1034	2	Q6N084	Q6N084 homo sapien
29	75	61.5	1255	2	Q6MZS0	Q6MZS0 homo sapien
30	75	61.5	1328	1	FINC PLEWA	Q91289 pleurodeles
31	75	61.5	2146	2	Q68CX6	Q68CX6 homo sapien

32 75 61.5 2193 2 Q6MZM7
33 67 54.9 2408 2 Q6JAN2
34 58 47.5 351 2 O02816
35 54 44.3 76 2 O46557
36 49.5 40.6 332 2 Q9MC76
37 49 40.2 156 2 Q70JQ4
38 49 40.2 170 2 Q8HSL8
39 48.5 39.8 197 2 Q7XWJ0
40 48 39.3 67 2 Q8MRE1
41 48 39.3 554 1 NTP3 TOBAC
42 48 39.3 680 2 Q7PXM0
43 47.5 38.9 416 2 Q63L42
44 47 38.5 96 2 Q7YTR6
45 47 38.5 178 1 CRB2_BIFLO

ALIGNMENTS

RESULT 1
Q8C6J7 PRELIMINARY; PRT; 296 AA.
AC Q8C6J7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330027109 product:fibronectin 1, full insert sequence.
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitanai T., Taishiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,

RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.",
RN Genome Res. 10:1757-1771(2000).
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirakata T., Hirozane T.,
RA Hori F., Imotani K., Iehii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tonaru A., Taya T., Yaeunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AK054456; BAC35784.1; -;
DR HSSP; Q96KP7; 1FBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR Pfam; PF00039; fnl; 5.
DR PRINTS; PR00012; FNTYPEI.
DR SMART; SM00058; FN1; 5.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 5.
SQ SEQUENCE 296 AA; 32617 MW; A0PF5F4809FB6439 CRC64;

Query Match 100.0%; Score 122; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGALVCTCYGSSR 21
Db 64 QWERTYLGALVCTCYGSSR 84

RESULT 2

ID Q72391 PRELIMINARY; PRT; 749 AA.
AC Q72391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloeker H., Boecher M., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -;
DR HSSP; Q96KP7; 1FBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fnl; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000395; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 9.

DR PROSITE; PS00023; FIBRONECTIN 2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 122; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 6.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGALVCTCYGSSR 21
Db 155 QWERTYLGALVCTCYGSSR 175

RESULT 3

ID Q6MZFA PRELIMINARY; PRT; 1103 AA.
AC Q6MZFA;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219 (Fragment)).
DE Name=DKFZp686K139; Synonyms=DKFZp686F219;
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Oeanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CA546200.1; -;
DR EMBL; BX640802; CA545885.1; -;
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fnl; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000395; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 9.
DR PROSITE; PS00023; FIBRONECTIN 2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;

Query Match 100.0%; Score 122; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 8.8e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGALVCTCYGSSR 21
Db 152 QWERTYLGALVCTCYGSSR 172

RESULT 4

Q68DF9
ID Q68DF9 PRELIMINARY; PRT; 2267 AA.
AC Q68DF9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.59216 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-6
Perfect score: 122
Sequence: 1 QWERTYLGNALVCTCYGSSR 21
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	2477	2	FN4428
2	118	96.7	2386	1	FNHU
3	117	95.9	2265	1	FNBO
4	114	93.4	2481	2	A43908
5	75	61.5	103	2	A49173
6	48	39.3	554	2	S22495
7	47	38.5	210	2	D89606
8	47	38.5	265	2	S34668
9	47	38.5	552	2	T10850
10	47	38.5	774	2	F96639
11	46	37.7	638	2	AG1422
12	46	37.7	655	1	A46688
13	45.5	37.3	354	1	XSBPL5
14	45	36.9	106	2	T49958
15	45	36.9	106	2	S60232
16	45	36.9	280	2	T41986
17	45	36.9	300	1	D34443
18	45	36.9	300	2	AD1988
19	45	36.9	335	2	T47274
20	45	36.9	436	2	H64641
21	45	36.9	439	2	A18173
22	44.5	36.5	98	2	C84323
23	44.5	36.5	409	2	T08928
24	44.5	36.5	1234	1	NBM5H
25	44	36.1	65	2	A10051
26	44	36.1	149	2	T31446
27	44	36.1	190	2	I51279
28	44	36.1	276	2	H83927
29	44	36.1	297	2	F75263

hypothetical prote
apolipoprotein H-r
complement factor
complement factor
heamin storage sys
hypothetical prote
hypothetical prote
T-cell receptor be
185K secretory pro
laminin - Hydra vu
conserved hypothet
hypothetical prote
198 pc receptor al
probable D-aminope
probable transamin
trp-asp repeat pro

ALIGNMENTS

RESULT 1

S14428

fibronectin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jul-2004

C:Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049

R:Hynes, R.O.

submitted to the EMBL Data Library, July 1989

A:Reference number: S14428

A:Accession: S14428

A:Molecule type: mRNA

A:Residues: 1-2477 <HYN>

A:Cross-references: UNIPROT:P04937; EMBL:X15906; NID:G56163; PIDN:CAA34020.1; PID:G56164

R:Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.

EMBO J. 6, 2573-2580, 1987

A:Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.

A:Reference number: S12455; MUID:88054951; PMID:2445560

A:Accession: S12455

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 609-1810, 'T', 1812-2283 <SCH>

A:Cross-references: EMBL:X15906

R:Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.

Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984

A:Title: A single rat fibronectin gene generates three different mRNAs by alternative sp

A:Reference number: A22319; MUID:84298097; PMID:6089177

A:Accession: A22319

A:Molecule type: DNA

A:Residues: 2052-2237 <TAM>

R:Palkenberg, C.; Engchild, J.J.; Thogersen, I.B.; Salvesen, G.; Akerstroem, B.

Biochem. J. 301, 745-751, 1994

A:Title: Isolation and characterization of fibronectin-alpha(1)-microglobulin complex in

A:Reference number: S46203; MUID:94330948; PMID:7519849

A:Accession: S46203

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1183-1192; 'GUN', 1268, 'P', 1270-1271, 'D', 1273, 'CF', 1276, 'PV', 1385-1399 <PAL>

R:Patel, R.S.; Odermatt, E.; Schwarzbauer, J.E.; Hynes, R.O.

EMBO J. 6, 2565-2572, 1987

A:Title: Organization of the fibronectin gene provides evidence for exon shuffling durin

A:Reference number: S00459; MUID:88054950; PMID:3119323

A:Accession: S00459

A:Molecule type: DNA

A:Residues: 1-139; 2382-2477 <PAT>

A:Cross-references: EMBL:X05831

A:Note: the authors translated the codon CCT for residues 51 and 94 as Ala

R:Schwarzbauer, J.E.; Tamkun, J.W.; Lemischka, I.R.; Hynes, R.O.

Cell 35, 421-431, 1983

A:Title: Three different fibronectin mRNAs arise by alternative splicing within the codi

A:Reference number: A27252; MUID:84082067; PMID:6317187

A:Accession: A27252

A:Molecule type: mRNA

A;Residues: 1586-1720,'T',1722,1813-2477 <SC2>
R;Odermatt, E.; Tankun, J.W.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 82, 6571-6575, 1985
A;Title: Repeating modular structure of the fibronectin gene: Relationship to protein structure
A;Reference number: 159049; MUID:86016741; PMID:3863113
A;Accession: 159049
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1722-1810 <RES>
A;Cross-references: GB:M11750; NID:g204164; PIDN:AAA41170.1; PID:g554437
C;Genetics:
F;Introns: 51/1; 94/1; 2416/3; 2454/3
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repeat homology
C;Keywords: alternative splicing; cell adhesion; collagen binding; disulfide bond; duplication
F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-2477/Product: fibronectin #status predicted <MAT>
F;53-88/Domain: fibronectin type I repeat homology <1F1>
F;98-136/Domain: fibronectin type I repeat homology <1F2>
F;142-180/Domain: fibronectin type I repeat homology <1F3>
F;187-226/Domain: fibronectin type I repeat homology <1F4>
F;232-271/Domain: fibronectin type I repeat homology <1F5>
F;308-342/Domain: fibronectin type I repeat homology <1F6>
F;360-401/Domain: fibronectin type II repeat homology <2F1>
F;420-461/Domain: fibronectin type II repeat homology <2F2>
F;470-508/Domain: fibronectin type I repeat homology <1F7>
F;518-558/Domain: fibronectin type I repeat homology <1F8>
F;561-599/Domain: fibronectin type I repeat homology <1F9>
F;609-692/Domain: fibronectin type III repeat homology <FN3A>
F;718-800/Domain: fibronectin type III repeat homology <FN3B>
F;809-890/Domain: fibronectin type III repeat homology <FN3C>
F;905-987/Domain: fibronectin type III repeat homology <FN3D>
F;995-1076/Domain: fibronectin type III repeat homology <FN3E>
F;1085-1164/Domain: fibronectin type III repeat homology <FN3F>
F;1172-1257/Domain: fibronectin type III repeat homology <FN3G>
F;1265-1348/Domain: fibronectin type III repeat homology <FN3H>
F;1356-1439/Domain: fibronectin type III repeat homology <FN3I>
F;1447-1529/Domain: fibronectin type III repeat homology <FN3J>
F;1537-1619/Domain: fibronectin type III repeat homology <FN3K>
F;1634-1616/Region: cell attachment (R-G-D) motif
F;1634-1713/Domain: fibronectin type III repeat homology <FN3L>
F;1721-1803/Domain: fibronectin type III repeat homology <FN3M>
F;1811-1893/Domain: fibronectin type III repeat homology <FN3N>
F;1903-1984/Domain: fibronectin type III repeat homology <FN3O>
F;1992-2074/Domain: fibronectin type III repeat homology <FN3P>
F;2181-2183/Region: cell attachment (R-G-D) motif
F;2193-2273/Domain: fibronectin type III repeat homology <FN3Q>
F;2296-2335/Domain: fibronectin type I repeat homology <1F10>
F;2341-2378/Domain: fibronectin type I repeat homology <1F11>
F;2385-2420/Domain: fibronectin type I repeat homology <1F12>
F;53-79,77-88,98-126,124-136,142-170,168-180,187-216,214-226,232-261,259-271,308-335,333-368,2366-2378,2385-2411,2409-2420/Disulfide bonds: #status predicted
F;2458/Disulfide bonds: interchain (to 2462) #status predicted
F;2462/Disulfide bonds: interchain (to 2458) #status predicted

Query Match 100.0%; Score 122; DB 2; Length 2477;
Best Local Similarity 100.0%; Pred. No. 7.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQWERTYLGNALVCTCYGGSR 21
|||||
Db 64 QQWERTYLGNALVCTCYGGSR 84
|||||

RESULT 2
FNHU
N;Alternate names: fibronectin
C;Species: Homo sapiens (man)
C;Date: 27-Nov-1985 #sequence_revision 31-Mar-1993 #text_change 09-Jul-2004
A;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22368; D.C.; Bowlus, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656
A;Accession: A26460
A;Molecule type: DNA
A;Residues: 1-49 <DEA>
A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA5337
R;Oldberg, A.; Ruoslahti, E.
J. Biol. Chem. 261, 2113-2116, 1986
A;Title: Evolution of the fibronectin gene.
A;Reference number: A26284; MUID:86111901; PMID:3003095
A;Accession: A26284
A;Molecule type: DNA
A;Residues: 1447-1540 <OLD>
A;Cross-references: GB:M12549; NID:g182688
A;Note: the authors translated the codon TTC for residue 1494 as Glu
R;Paoletti, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
Nucleic Acids Res. 16, 3545-3557, 1988
A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
A;Reference number: S00848; MUID:88233940; PMID:3375063
A;Accession: S03917
A;Molecule type: DNA
A;Residues: 1594-1767,'V',1769-1783 <PAO>
A;Cross-references: EMBL:X07710; NID:g31402
A;Note: the authors translated the codon AAC for residue 1631 as Asp
R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
PBBS Lett. 207, 287-291, 1986
A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:
A;Reference number: A24854; MUID:87030929; PMID:3770201
A;Accession: A24854
A;Molecule type: DNA
A;Residues: 1992-2147 <VIB>
A;Cross-references: GB:X04530; NID:g31436
R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.
PBBS Lett. 207, 145-148, 1986
A;Title: Human fibronectin is synthesized as a pre-propolypeptide.
A;Reference number: A24476; MUID:87030890; PMID:3770189
A;Accession: A24476
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-14,'Q',16-38 <GUT>
R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A;Title: Primary structure of human fibronectin: differential splicing may generate at
A;Reference number: A91008; MUID:85284965; PMID:2992939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080;2112-2386 <KOR>
A;Cross-references: GB:X02761
R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 973-2080;2112-2386 <KO2>
A;Cross-references: GB:X00739
R;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
A;Cross-references: GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with r
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386 <BER>
A;Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
R;Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 7.1098 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QQWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	231	9	US-09-925-302-548
2	122	100.0	231	10	Sequence 548, App
3	122	100.0	463	15	US-09-925-302-548
4	122	100.0	463	16	Sequence 548, App
5	122	100.0	642	16	US-10-144-194A-52
6	122	100.0	642	17	US-10-491-566-52
7	122	100.0	657	16	US-10-741-601-354
8	122	100.0	657	17	US-10-741-601-1066
9	122	100.0	657	16	US-10-741-601-359
10	122	100.0	984	16	US-10-741-600-1072
11	122	100.0	984	17	US-10-741-601-356
12	122	100.0	1173	18	US-10-741-600-1069
13	122	100.0	1173	18	US-10-450-763-52634

12	122	100.0	2220	15	US-10-236-392-4	Sequence 4, Appl
13	122	100.0	2296	16	US-10-741-601-363	Sequence 363, App
14	122	100.0	2296	17	US-10-741-600-1075	Sequence 1075, App
15	122	100.0	2320	14	US-10-279-733-8	Sequence 8, Appl
16	122	100.0	2355	15	US-10-144-194A-104	Sequence 104, App
17	122	100.0	2355	15	US-10-360-101-235	Sequence 235, App
18	122	100.0	2355	15	US-10-447-161-3	Sequence 3, Appl
19	122	100.0	2355	16	US-10-734-564-94	Sequence 94, Appl
20	122	100.0	2355	16	US-10-741-601-366	Sequence 366, App
21	122	100.0	2355	16	US-10-741-601-366	Sequence 366, App
22	122	100.0	2355	16	US-10-491-566-104	Sequence 104, App
23	122	100.0	2355	17	US-10-741-600-1067	Sequence 1067, App
24	122	100.0	2355	17	US-10-741-600-1078	Sequence 1078, App
25	122	100.0	2355	17	US-10-852-335A-147	Sequence 147, App
26	122	100.0	2355	18	US-10-287-436A-436	Sequence 436, App
27	122	100.0	2355	18	US-10-287-436A-1137	Sequence 1137, App
28	122	100.0	2355	20	US-11-040-130-28	Sequence 28, Appl
29	122	100.0	2386	10	US-09-961-403-1	Sequence 1, Appl
30	122	100.0	2386	16	US-10-741-601-360	Sequence 360, App
31	122	100.0	2386	16	US-10-788-792-206	Sequence 206, App
32	122	100.0	2386	16	US-10-618-281-32	Sequence 32, Appl
33	122	100.0	2386	17	US-10-741-600-1071	Sequence 1071, App
34	122	100.0	2386	17	US-10-868-577A-59	Sequence 59, Appl
35	122	100.0	2386	18	US-10-485-758-4	Sequence 4, Appl
36	122	100.0	2386	18	US-10-485-758-9	Sequence 9, Appl
37	118	96.7	259	10	US-09-940-235-4	Sequence 4, Appl
38	118	96.7	2320	15	US-10-236-392-2	Sequence 2, Appl
39	118	96.7	2328	14	US-10-171-311-64	Sequence 64, Appl
40	118	96.7	2328	15	US-10-236-031B-70	Sequence 70, Appl
41	118	96.7	2328	15	US-10-374-979-98	Sequence 98, Appl
42	118	96.7	2328	15	US-10-182-936A-98	Sequence 98, Appl
43	118	96.7	2328	16	US-10-477-238A-677	Sequence 677, App
44	118	96.7	2328	16	US-10-680-287A-677	Sequence 677, App
45	118	96.7	2328	17	US-10-477-173-677	Sequence 677, App

ALIGNMENTS

RESULT 1
US-09-925-302-548
; Sequence 548, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P4104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 548
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-302-548

```

RESULT 6
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al
; TITLE OF INVENTION: GENETIC POLYMORPHISMS
; TITLE OF INVENTION: MYOCARDIAL INFARCTION
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10741,600

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.22353 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QWERTYLGALVCTCYGSR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	2386	4	US-09-961-403-1
2	118	96.7	2231	1	US-08-153-799-16
3	118	96.7	2324	1	US-08-283-857-1
4	118	96.7	2324	5	PCT-US95-09819-1
5	118	96.7	2327	6	5455158-1
6	118	96.7	2327	6	5455158-1
7	118	96.7	2386	2	US-09-016-366A-12
8	118	96.7	2446	2	US-08-551-356-2
9	118	96.7	2446	5	PCT-US93-12687-2
10	75	61.5	188	1	US-08-142-449B-14
11	75	61.5	422	2	US-08-836-854-12
12	75	61.5	474	2	US-08-836-854-9
13	54	44.3	42	1	US-08-340-812-7
14	54	44.3	42	1	US-08-459-064B-7
15	54	44.3	42	1	US-08-460-421A-7
16	54	44.3	42	5	PCT-US93-00909-7
17	50.5	41.4	48	2	US-08-982-597A-20
18	50.5	41.4	48	3	US-09-136-218-20
19	47	38.5	668	4	US-09-248-796A-18675
20	46	37.7	276	4	US-09-270-767-43364
21	46	37.7	655	1	US-08-148-910-12
22	46	37.7	655	1	US-08-448-937A-12
23	45	36.9	37	2	US-08-982-597A-17
24	45	36.9	37	3	US-09-136-218-17
25	45	36.9	336	4	US-09-252-991A-20404
26	44	36.1	123	3	US-08-793-450-4
27	44	36.1	353	4	US-09-543-681A-5307

Sequence 14, Appl
Sequence 14, Appl
Sequence 13133, A
Sequence 18, Appl
Sequence 18, Appl
Sequence 24, Appl
Sequence 7031, Ap
Sequence 2, Appli
Sequence 5150, Ap
Sequence 21, Appl
Sequence 21, Appl
Sequence 5158, Ap
Sequence 1560, Ap
Sequence 7975, Ap
Sequence 12, Appl
Sequence 12, Appl
Sequence 12, Appl

28 43 35.2 227 1 US-07-869-933-14
29 43 35.2 227 3 US-09-103-663-14
30 43 35.2 544 4 US-09-489-039A-13133
31 43 35.2 587 2 US-08-871-266B-18
32 43 35.2 587 2 US-09-018-864A-18
33 43 35.2 587 3 US-08-871-267B-24
34 43 35.2 587 3 US-09-618-419-24
35 43 35.2 593 4 US-09-949-016-7031
36 43 35.2 2321 4 US-09-230-652-2
37 42.5 34.8 232 4 US-09-328-352-5150
38 42 34.4 43 2 US-08-982-597A-21
39 42 34.4 43 3 US-09-136-218-21
40 42 34.4 56 4 US-09-513-999C-5158
41 42 34.4 143 4 US-09-732-210-1560
42 42 34.4 143 4 US-09-513-999C-7975
43 42 34.4 164 4 US-09-907-794A-12
44 42 34.4 164 4 US-09-905-125A-12
45 42 34.4 164 4 US-09-902-775A-12

ALIGNMENTS

RESULT 1
US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961.403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

Query Match 100.0%; Score 122; DB 4; Length 2386;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGALVCTCYGSR 21
Db 63 QWERTYLGALVCTCYGSR 83

RESULT 2
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5768883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/153,799
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847975
FILING DATE: 06-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8909916.2
FILING DATE: 29-APR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB90/00650
FILING DATE: 26-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/775952
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Swope, R Hain
REGISTRATION NUMBER: 24864
REFERENCE/DOCKET NUMBER: 92H832
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 665 2400
TELEFAX: (908) 771 6159
TELEX: 219484
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 2231 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..2231
OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match 96.7%; Score 118; DB 1; Length 2231;
Best Local Similarity 95.2%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QQWERTYLGNALVCTCYGGSR 21
Db 32 QQWERTYLGNVLVCTCYGGSR 52

RESULT 3
US-08-283-857-1
Sequence 1, Application US/08283857
Patent No. 5792742
GENERAL INFORMATION:
APPLICANT: GOLD, Leslie I.
APPLICANT: ROSTAGNO, Agueda A.
APPLICANT: BARON, Martin
APPLICANT: CAMPBELL, Iain D.
APPLICANT: WILLIAMS, Michael, J.
TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/283,857
FILING DATE: 01-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/714,134
FILING DATE: 14-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-283-857-1

Query Match 96.7%; Score 118; DB 1; Length 2324;
Best Local Similarity 95.2%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QQWERTYLGNALVCTCYGGSR 21
Db 32 QQWERTYLGNVLVCTCYGGSR 52

RESULT 4
PCT-US95-09819-1
Sequence 1, Application PC/TUS9509819
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/09819
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/283,857
FILING DATE: 01-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: GOLD-1A PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2324 amino acids

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.76863 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QQWERTYLGALVCTCYGSSR 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep_16Dec04.*

- 1: Genesep1980s.*
- 2: Genesep1990s.*
- 3: Genesep2000s.*
- 4: Genesep2001s.*
- 5: Genesep2002s.*
- 6: Genesep2003as.*
- 7: Genesep2003bs.*
- 8: Genesep2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	21	2	AAY28903 MSF 1-alp
2	122	100.0	231	3	AAB58210 Lung canc
3	122	100.0	463	6	ABR58303 BCU0770 p
4	122	100.0	642	8	ADQ39403 Human myo
5	122	100.0	642	8	ADR67316 Human bla
6	122	100.0	642	8	ADSL17489 Amino aci
7	122	100.0	642	8	ADR97658 Human fib
8	122	100.0	657	8	ADQ39409 Human myo
9	122	100.0	660	2	AAY28901 Human mig
10	122	100.0	984	8	ADQ39406 Human myo
11	122	100.0	1173	4	ABG22275 Novel hum
12	122	100.0	1223	8	ADP30365 Human sec
13	122	100.0	2182	8	ADR66462 Human pro
14	122	100.0	2182	8	ADR66120 Human pro
15	122	100.0	2220	6	ABO01289 Human pro
16	122	100.0	2220	8	ADN95950 Human NOV
17	122	100.0	2265	4	RAM38647 Human pol
18	122	100.0	2266	6	ABR40124 Human cel
19	122	100.0	2296	8	ADQ39412 Human myo
20	122	100.0	2320	6	RAE37107 Human fib
21	122	100.0	2320	6	ABR42588 Human fib
22	122	100.0	2330	4	RAM38646 Human pol
23	122	100.0	2355	4	RAM38649 Human pol
24	122	100.0	2355	6	ABR58335 NM_00202
25	122	100.0	2355	7	ADP65196 Human fib

26	122	100.0	2355	8	ADG89560 Human fib
27	122	100.0	2355	8	ADL92160 Fibronect
28	122	100.0	2355	8	ADO55175 Protein #
29	122	100.0	2355	8	ADQ26085 Fibronect
30	122	100.0	2355	8	ADQ29668 Human col
31	122	100.0	2355	8	ADQ39415 Human myo
32	122	100.0	2355	8	ADQ39404 Human myo
33	122	100.0	2355	8	ADR67315 Human bla
34	122	100.0	2386	4	AAM38648 Human pol
35	122	100.0	2386	5	AAO17353 Human fib
36	122	100.0	2386	6	ABR81866 Human fib
37	122	100.0	2386	7	ADD18770 Human dis
38	122	100.0	2386	7	ADB63324 Human Pro
39	122	100.0	2386	8	ADO55174 Protein #
40	122	100.0	2386	8	ADQ39408 Human myo
41	122	100.0	2386	8	ADR99200 Fibronect
42	122	100.0	2446	3	AAB50377 Human fib
43	122	100.0	2447	4	AAM40434 Human pol
44	122	100.0	2447	4	AAM40435 Human pol
45	122	100.0	2447	4	AAM40433 Human pol

ALIGNMENTS

RESULT 1

AAY28903

ID AAY28903 standard; peptide; 21 AA.

AC AAY28903;

DT 21-SEP-1999 (first entry)

DE MSF 1-alpha peptide epitope.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.
XX Synthetic.
OS Homo sapiens.
XX WO9931233-A1.
PD 24-JUN-1999.
PF 15-DEC-1998; 98WO-GB003766.
PR 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.
XX Claim 18; Page 57; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAY28902-906 represent peptide epitopes of MSF against which monoclonal antibodies that are specific to MSF and do not cross-react with fibronectin are raised

SQ Sequence 21 AA;

Query Match 100.0%; Score 122; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSR 21
 Db 1 QWERTYLGNALVCTCYGSR 21

RESULT 2
 AAB58210
 ID AAB58210 standard; protein; 231 AA.
 XX AC AAB58210;
 XX
 DT 14-MAR-2001 (first entry)
 DE Lung cancer associated polypeptide sequence SEQ ID 548.
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnerary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX OS Homo sapiens.
 XX PN WO200055180-A2.
 XX PD 21-SEP-2000.
 XX PF 08-MAR-2000; 2000WO-US005918.
 XX PR 12-MAR-1999; 99US-0124270P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI (ROSE/) ROSEN C A.
 XX PI Ruben SM;
 XX WPI; 2000-587514/55.
 DR N-PSDB; AAF18086.
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer.
 XX Claim 11; Page 1041-1042; 1425pp; English.
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the protein
 CC or polynucleotide sequences. The lung cancer associated polynucleotide
 CC sequences may be used for detection of lung cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The proteins may be used to treat disorders such as
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
 CC cardiovascular, renal, and proliferative disorders. The proteins may also
 CC be used in the treatment of wounds and infectious diseases.
 CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
 CC used in the course of the invention for the identification and
 CC characterization of the polynucleotide and protein sequences
 XX Sequence 231 AA;
 SQ

Query Match 100.0%; Score 122; DB 3; Length 231;
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSR 21
 Db 1 QWERTYLGNALVCTCYGSR 21

RESULT 4
 AAB58210
 ID AAB58210 standard; protein; 231 AA.
 XX AC AAB58210;
 XX
 DT 14-MAR-2001 (first entry)
 DE Lung cancer associated polypeptide sequence SEQ ID 548.
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardioactive; immunomodulatory; muscular active; vulnerary;
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX OS Homo sapiens.
 XX PN WO200055180-A2.
 XX PD 21-SEP-2000.
 XX PF 08-MAR-2000; 2000WO-US005918.
 XX PR 12-MAR-1999; 99US-0124270P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI (ROSE/) ROSEN C A.
 XX PI Ruben SM;
 XX WPI; 2000-587514/55.
 DR N-PSDB; AAF18086.
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer.
 XX Claim 11; Page 1041-1042; 1425pp; English.
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the protein
 CC or polynucleotide sequences. The lung cancer associated polynucleotide
 CC sequences may be used for detection of lung cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The proteins may be used to treat disorders such as
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
 CC cardiovascular, renal, and proliferative disorders. The proteins may also
 CC be used in the treatment of wounds and infectious diseases.
 CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
 CC used in the course of the invention for the identification and
 CC characterization of the polynucleotide and protein sequences
 XX Sequence 231 AA;
 SQ

Query Match 100.0%; Score 122; DB 3; Length 231;
 Best Local Similarity 100.0%; Pred. No. 1.8e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 92 QWERTYLGNALVCTCYGSR 112

RESULT 3
 ABR58303
 ID ABR58303 standard; protein; 463 AA.
 XX AC ABR58303;
 XX DT 08-JUL-2003 (first entry)
 XX DB BCU0770 protein #SEQ ID 52.
 KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 KW drug discovery; clinical medicine; forensic medicine; chromosome 2q35.
 XX OS Homo sapiens.
 XX PN WO2003029421-A2.
 XX PD 10-APR-2003.
 XX PF 02-OCT-2002; 2002WO-US031287.
 XX PR 03-OCT-2001; 2001US-0326526P.
 XX PR 14-MAY-2002; 2002US-00144194.
 XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.
 XX PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
 XX WPI; 2003-381623/36.
 DR N-PSDB; ACC72037.
 PT New isolated human differentially-regulated breast cancer polynucleotide
 PT and polypeptide, useful for diagnosing, staging, prognosticating,
 PT preventing and/or treating diseases and conditions relating to breast
 PT cancer.
 XX Claim 6; SEQ ID NO 52; 127pp + Sequence Listing; English.
 PS The invention relates to isolated polynucleotides which are
 CC differentially-regulated in breast cancer. The methods and compositions
 CC of the present invention are useful for detecting, diagnosing, staging,
 CC monitoring, prognosticating, preventing and/or treating diseases and
 CC conditions relating to breast cancer, and may be used in gene therapy or
 CC antisense therapy. They can also be used in research, drug discovery,
 CC clinical medicine and forensic medicine. Sequences given in records
 CC ABR58278-ABR58323 and ABR58346-ABR58362 represent polypeptides encoded by
 CC polynucleotides of the invention that are differentially-regulated in
 CC breast cancer. NOTE: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 463 AA;
 SQ

Query Match 100.0%; Score 122; DB 6; Length 463;
 Best Local Similarity 100.0%; Pred. No. 3.5e-08;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QWERTYLGNALVCTCYGSR 21
 Db 99 QWERTYLGNALVCTCYGSR 119

RESULT 4
 ADQ39403
 ID ADQ39403 standard; protein; 642 AA.
 XX AC ADQ39403;
 XX DT 18-NOV-2004 (first entry)
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 436.745 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-6
Perfect score: 122
Sequence: 1 QQWERTYLGALVCTCYGSSR 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2 1/USPTO_spool_b/US09581651/runat_07112005_092223_28789/app_query.fasta.1.1834
-DB=EST -QFMT=fascap -SUPFIX=p2n.rst -MINMATCH=0.1 -LOOPCD=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCFALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651_QCGN 1.1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: gb_est1.*

2: gb_est2.*

3: gb_hcc.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_g881.*

9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

C 1	122	100.0	239	2	BF358738 CM2-ET001
2	122	100.0	292	4	BF988947 IL5-GN017
3	122	100.0	305	2	AW068290 cn20c07.Y
4	122	100.0	317	9	CU212775 G050E08 G
5	122	100.0	321	2	AW068245 cn23g07.Y
6	122	100.0	343	2	B8817493 QV4-BN026
7	122	100.0	349	6	CB704303 AMGNNUC:C
C 8	122	100.0	378	2	BF751799 RC3-BN042
9	122	100.0	407	9	CG536337 OST124189

C 10	122	100.0	427	4	BF988908
11	122	100.0	435	1	AU140826
C 12	122	100.0	446	2	BF329480
13	122	100.0	465	7	CF166902
14	122	100.0	473	7	CF166406
15	122	100.0	480	4	BF993488
16	122	100.0	481	7	CF896895
C 17	122	100.0	487	4	BF993485
C 18	122	100.0	491	4	BF993489
19	122	100.0	494	5	BQ366579
20	122	100.0	494	7	CN332336
21	122	100.0	514	1	AL710256
22	122	100.0	516	5	BX509989
23	122	100.0	521	5	BX646566
24	122	100.0	525	5	BX473822
25	122	100.0	528	5	BX470663
26	122	100.0	530	1	AL702029
27	122	100.0	530	5	BX473863
28	122	100.0	532	5	BX473866
29	122	100.0	535	1	AL699665
30	122	100.0	536	5	BX474091
31	122	100.0	537	5	BX473795
32	122	100.0	539	5	BX473834
33	122	100.0	540	5	BX473642
34	122	100.0	541	5	BX504656
35	122	100.0	546	5	BX473798
36	122	100.0	548	5	BX473892
37	122	100.0	548	5	BX473934
38	122	100.0	548	7	CN332331
39	122	100.0	554	5	BX473713
40	122	100.0	561	5	BX473828
41	122	100.0	562	5	BX470610
42	122	100.0	565	1	AL702774
43	122	100.0	565	1	AL702809
44	122	100.0	565	1	AL712567
45	122	100.0	565	5	BX508450

ALIGNMENTS

RESULT 1
BF358738/c

LOCUS BF358738 239 bp mRNA linear EST 22-NOV-2000
DEFINITION CM2-ET0016-310500-206-f05 ET0016 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF358738
VERSION BF358738.1 GI:11317810
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brundstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?c1=CM2&t2=CM2-ET0016-310500-206-f05&t3=2000-05-31&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 239.

FEATURES

Location/Qualifiers
 1..239
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="ET0016"

/note="Organ: lung tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 1.13e-08 Length: 239
 Score: 122.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-581-651D-6 (1-21) x BF358738 (1-239)

Qy 1 GlnGlnTrpGluArgThrTyLeuGlyAsnAlaLeuValCysThrCysTyGlyGlySer 20
 |||||
 Db 190 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGAGGAAGC 131

Qy 21 Arg 21
 |||
 Db 130 CGA 128

RESULT 2

BF988947
 LOCUS IL5-GN0179-181000-176-c08 GN0179 Homo sapiens cDNA, mRNA sequence. EST 23-JAN-2001

ACCESSION BF988947

VERSION BF988947.1 GI:12395272

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?c1=IL5&t2=IL5-GN0179-310500-206-f05&t3=2000-05-31&t4=1>)

181000-176-c08&t3=2000-10-18&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 291.

FEATURES

source

Location/Qualifiers
 1..292
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="GN0179"

/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 1.43e-08 Length: 292
 Score: 122.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-581-651D-6 (1-21) x BF988947 (1-292)

Qy 1 GlnGlnTrpGluArgThrTyLeuGlyAsnAlaLeuValCysThrCysTyGlyGlySer 20
 |||||
 Db 193 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGAGGAAGC 252

Qy 21 Arg 21
 |||

Db 253 CGA 255

RESULT 3

AW068290

LOCUS

DEFINITION

cn20c07.y1 Normal Human Trabecular Bone Cells Homo sapiens CDNA

clone NHTBC_cn20c07 random, mRNA sequence.

ACCESSION AW068290

VERSION AW068290.1 GI:6023288

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G., Yang,L.M.,

Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,

Robey,P.G., Hotchkiss,R.N. and Franccomano,C.A.

SGAP: The Skeletal Genome Anatomy Project

Unpublished (1997)

JOURNAL

COMMENT

Contact: Libin Jia

Medical Genetics Branch

National Human Genome Research Institute

10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA

Tel: 301-402-4877

Fax: 301-496-7157

Email: libin@helix.nih.gov

DNA Sequencing and analyses by National Institutes of Health

Intramural Sequencing Center (NISC).

Plate: 20 row: c column: 07

Seq primer: M13RP1 reverse primer (ABI).

Location/Qualifiers

1..305

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="NHTBC_cn20c07"

/sex="Female"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 107.827 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-6
Perfect score: 122
Sequence: 1 QQWERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=Published Applications NA -Qfmt=fastap -Suffix=p2n.rnpb -MINMATCH=0.1
-TRANSPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -THR SCORE=pct -THR MAX=100
-MAXLEN=2000000000 -USER=US09581651 @CGN 1 1 1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=3 -NO MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT=-DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	122	100.0	407	9	US-09-728-445-799 Sequence 799, App
2	122	100.0	407	24	US-10-964-549-799 Sequence 799, App
3	122	100.0	867	9	US-09-925-302-105 Sequence 105, App
4	122	100.0	867	10	US-09-925-302-105 Sequence 105, App
5	122	100.0	2127	17	US-10-210-120-49 Sequence 49, Appl
6	122	100.0	2127	22	US-10-956-157-4288 Sequence 4288, Ap
7	122	100.0	2127	24	US-10-909-035-49 Sequence 49, Appl
8	122	100.0	2443	20	US-10-741-601-70 Sequence 70, Appl
9	122	100.0	2443	22	US-10-741-600-238 Sequence 238, App
10	122	100.0	2488	20	US-10-741-601-75 Sequence 75, Appl
11	122	100.0	3522	24	US-10-450-763-22266 Sequence 2266, A
12	122	100.0	4295	18	US-10-144-194A-51 Sequence 51, Appl
13	122	100.0	4295	21	US-10-491-566-51 Sequence 51, Appl
14	122	100.0	6510	20	US-10-741-601-72 Sequence 72, Appl
15	122	100.0	6510	22	US-10-741-600-241 Sequence 241, App
16	122	100.0	7361	19	US-10-236-392-3 Sequence 3, Appl1
17	122	100.0	7361	19	US-10-236-392-3 Sequence 3, Appl1
18	122	100.0	7677	22	US-10-956-157-4995 Sequence 4995, Ap
19	122	100.0	7795	16	US-10-084-817-2 Sequence 2, Appl1
20	122	100.0	7823	20	US-10-741-601-77 Sequence 77, Appl
21	122	100.0	7823	22	US-10-741-600-245 Sequence 245, App
22	122	100.0	7848	20	US-10-741-601-78 Sequence 246, App
23	122	100.0	7848	22	US-10-741-600-246 Sequence 246, App
24	122	100.0	7867	14	US-10-098-841-6 Sequence 6, Appl1
25	122	100.0	7935	20	US-10-741-601-74 Sequence 74, Appl
26	122	100.0	7935	22	US-10-741-600-240 Sequence 240, App
27	122	100.0	7959	20	US-10-741-601-81 Sequence 81, Appl
28	122	100.0	7959	22	US-10-741-600-249 Sequence 249, App
29	122	100.0	8013	20	US-10-741-601-71 Sequence 71, Appl
30	122	100.0	8013	22	US-10-741-600-242 Sequence 242, App
31	122	100.0	8027	18	US-10-447-161-8 Sequence 8, Appl1
32	122	100.0	8027	20	US-10-734-564-27 Sequence 27, Appl
33	122	100.0	8027	22	US-10-852-335A-53 Sequence 53, Appl
34	122	100.0	8027	24	US-10-287-436A-81 Sequence 121, App
35	122	100.0	8044	17	US-10-240-965-121 Sequence 121, App
36	122	100.0	8044	24	US-10-765-700-135 Sequence 135, App
37	122	100.0	8062	14	US-10-098-841-5 Sequence 5, Appl1
38	122	100.0	8137	14	US-10-098-841-8 Sequence 8, Appl1
39	122	100.0	8155	20	US-10-741-601-79 Sequence 79, Appl
40	122	100.0	8155	22	US-10-741-600-247 Sequence 247, App
41	122	100.0	8226	20	US-10-741-601-69 Sequence 69, Appl
42	122	100.0	8226	22	US-10-741-600-237 Sequence 237, App
43	122	100.0	8230	14	US-10-098-841-7 Sequence 7, Appl1
44	122	100.0	8278	20	US-10-741-601-82 Sequence 82, Appl
45	122	100.0	8278	22	US-10-741-600-250 Sequence 250, App

ALIGNMENTS

RESULT 1

US-09-728-445-799
; Sequence 799, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020102543A1e1 Mutated Mammalian Cells and
; TITLE OF INVENTION: Animals
; FILE REFERENCE: LEX-0102-USA
; CURRENT APPLICATION NUMBER: US/09/728,445
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358

```
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 799
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(407)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-445-799

Alignment Scores:
Pred. No.: 1,36e-10 Length: 407
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-728-445-799 (1-407)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 20
Db 283 CAGCAGTGGGAACGGACCTACCTAGGCAACGCCCTGGTTGTACCTGCTATGGAGGAAGC 342
Qy 21 Arg 21
Db 343 CGG 345

RESULT 2
US-10-964-549-799
; Sequence 799, Application US/10964549
; Publication No. US20050186677A1
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: Novel Mutated Mammalian Cells and Animals
; FILE REFERENCE: LEX-0286-USA
; CURRENT APPLICATION NUMBER: US/10/964,549
; CURRENT FILING DATE: 2004-10-13
; PRIOR APPLICATION NUMBER: US/09/750,456
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 09/728,445
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,358
; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 891
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 799
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(407)
; OTHER INFORMATION: n = A,T,C or G
US-10-964-549-799

Alignment Scores:
Pred. No.: 1,36e-10 Length: 407
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-581-651D-6 (1-21) x US-10-964-549-799 (1-407)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 20
Db 283 CAGCAGTGGGAACGGACCTACCTAGGCAACGCCCTGGTTGTACCTGCTATGGAGGAAGC 342
Qy 21 Arg 21
Db 343 CGG 345

RESULT 3
US-09-925-302-105
; Sequence 105, Application US/09925302
; Patent No. US2002004941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-105

Alignment Scores:
Pred. No.: 3,29e-10 Length: 867
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-925-302-105 (1-867)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 20
Db 448 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGAGGAAGC 507
Qy 21 Arg 21
Db 508 CGA 510

RESULT 4
US-09-925-302-105
; Sequence 105, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-105

Alignment Scores:
Pred. No.: 3,29e-10 Length: 867
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
DB: 9 Gaps: 0
```

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	122	100.0	8044	4	US-09-566-921-135	Sequence 135, App
2	118	96.7	7679	4	US-09-220-112-38	Sequence 38, Appl
3	118	96.7	7680	4	US-09-023-655-1289	Sequence 1289, Ap
4	118	96.7	7680	5	PT-US95-09819-6	Sequence 6, Appl
5	118	96.7	7705	1	US-08-259-569-16	Sequence 16, Appl
6	118	96.7	7705	2	US-08-826-885-16	Sequence 16, Appl
7	118	96.7	7705	6	5455158-2	Patent No. 5455158
8	118	96.7	7705	6	5455158-2	Patent No. 5455158
9	118	96.7	7803	2	US-08-551-356-1	Sequence 1, Appl
10	118	96.7	7803	5	PT-US93-12687-1	Sequence 1, Appl
c 11	75	61.5	241	4	US-09-389-631-347	Sequence 347, App
c 12	75	61.5	241	4	US-09-620-405B-347	Sequence 347, App

Qy 21 Arg 21
Db 515 CGA 517

RESULT 2
US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:
Pred. No.: 1.43e-08 Length: 7679
Score: 118.00 Matches: 20
Percent Similarity: 95.24% Conservative: 0
Best Local Similarity: 95.24% Mismatches: 1
Query Match: 96.72% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-220-132-38 (1-7679)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 107 CAACAGTGGAGCGGACCTACCTAGTAATGTGTGTTGTACTTGTATTGAGGAAGC 166

Qy 21 Arg 21
Db 167 CGA 169

RESULT 3
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:

PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528

PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528

PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g31396
US-09-023-655-1289

Alignment Scores:
Pred. No.: 1.43e-08 Length: 7680
Score: 118.00 Matches: 20
Percent Similarity: 95.24% Conservative: 0
Best Local Similarity: 95.24% Mismatches: 1
Query Match: 96.72% Indels: 0
DB: 4 Gaps: 0

US-09-581-651D-6 (1-21) x US-09-023-655-1289 (1-7680)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 107 CAACAGTGGAGCGGACCTACCTAGTAATGTGTGTTGTACTTGTATTGAGGAAGC 166

Qy 21 Arg 21
Db 167 CGA 169

RESULT 4
PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 68.8471 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122

Sequence: 1 QQMERTYLGALVCTCYGGR 21

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool.b/US09581651/runat_07112005_092222_28772/app_query.fasta_1.1834
-DB=N_Geneseq_QFT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651.QCGN 1 1 1052 @runat_07112005_092222_28772 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	100.0	407	6	ABs69736 Novel mur
2	122	100.0	867	3	Aaf18086 Lung canc
3	122	100.0	1929	13	Adg17488 Nucleotid
4	122	100.0	1929	13	Adr97657 Human fib
5	122	100.0	2127	10	Add18477 Human pro

6	122	100.0	2147	2	AAX81299	Aax81299 Human mig
7	122	100.0	2358	13	ADr67201	Adr67201 Human bla
8	122	100.0	2443	13	ADQ38575	Adq38575 Human SNP
9	122	100.0	2488	13	ADQ38581	Adq38581 Human SNP
10	122	100.0	3522	5	AAS86462	Aas86462 DNA encod
11	122	100.0	3669	12	ADP29134	Adp29134 Human sec
12	122	100.0	4295	8	ACC72037	Acc72037 BC00770 g
13	122	100.0	6510	13	ADQ38578	Adq38578 Human SNP
14	122	100.0	6816	10	ABX74443	Abx74443 Human CDN
15	122	100.0	7242	13	ADQ66637	Adq66637 Human pro
16	122	100.0	7242	13	ADQ65953	Adq65953 Human pro
17	122	100.0	7361	9	ACD06170	AcD06170 Human CDN
18	122	100.0	7550	8	ACC00412	Acc00412 Human cel
19	122	100.0	7677	13	ACN37820	Acn37820 Tumour-as
20	122	100.0	7795	10	ADJ56196	Adj56196 Zebrafish
21	122	100.0	7823	13	ADQ38582	Adq38582 Human SNP
22	122	100.0	7848	13	ADQ38583	Adq38583 Human SNP
23	122	100.0	7867	4	AAI57803	Aai57803 Human pol
24	122	100.0	7935	13	ADQ38577	Adq38577 Human SNP
25	122	100.0	7959	13	ADQ38586	Adq38586 Human SNP
26	122	100.0	8013	13	ADQ38579	Adq38579 Human SNP
27	122	100.0	8027	11	ADP64998	Adp64998 Human fib
28	122	100.0	8027	12	ADQ89565	Adq89565 Human fib
29	122	100.0	8027	12	ADQ29601	Adq29601 Human col
30	122	100.0	8027	13	ADP67200	Adp67200 Human bla
31	122	100.0	8039	3	AAC89889	Aac89889 Human FN
32	122	100.0	8044	6	AAS94866	Aas94866 Human DNA
33	122	100.0	8044	12	ADI61767	Adi61767 Human CDN
34	122	100.0	8062	4	AAI57802	Aai57802 Human pol
35	122	100.0	8066	4	AAI59591	Aai59591 Human pol
36	122	100.0	8066	4	AAI59589	Aai59589 Human pol
37	122	100.0	8066	4	AAI59588	Aai59588 Human pol
38	122	100.0	8066	4	AAI59590	Aai59590 Human pol
39	122	100.0	8137	4	AAI57805	Aai57805 Human pol
40	122	100.0	8155	13	ADQ38584	Adq38584 Human SNP
41	122	100.0	8226	13	ADQ38574	Adq38574 Human SNP
42	122	100.0	8230	4	AAI57804	Aai57804 Human pol
43	122	100.0	8278	13	ADQ38587	Adq38587 Human SNP
44	122	100.0	8329	6	ABK63824	Abk63824 Rat seque
45	122	100.0	8329	10	ADB58331	Adb58331 Toxicity-

ALIGNMENTS

RESULT 1

ABS69736
ID ABS69736 standard; DNA; 407 BP.

XX AC ABS69736;

XX DT 21-NOV-2002 (first entry)

XX DE Novel murine polynucleotide isolated using gene trap technology #799.

XX KW Mouse; gene trapped sequence; GTS; functional genomic analysis;
KW phage display system; gene chip; temporal gene expression;
KW tissue specific gene expression; antisense inhibition; gene targeting;
KW development disorder; cell differentiation disorder; aging; cancer;
KW autoimmune disease; lupus; inflammatory disorder; skin disorder;
KW degenerative disorder; ds.

XX OS Mus musculus.

XX PN US2002102543-A1.

XX PD 01-AUG-2002.

XX PF 30-NOV-2000; 2000US-00728445.

XX PR 01-DEC-1999; 99US-0168358P.

XX XX (FRIE/) FRIEDRICH G.

XX FA (ZAMB/) ZAMBROWICZ B.

PA (SAND/) SANDS A T.
XX Friedrich G, Zambrowicz B, Sands AT;
XX WPI; 2002-690598/74.
XX Novel murine polynucleotides that individually identify novel genes into
PT which a retroviral gene trap vector has integrated, useful in genomic
PT analysis and in discovery, development of therapeutic and diagnostic
XX agents.
XX Claim 1; Page 264; 296pp; English.
XX
XX The invention describes an isolated murine polynucleotide (I) comprising
CC a contiguous stretch of at least 60 nucleotides of one of 265-677
CC nucleotide 891 ONNIBANK gene trapped sequences (GTSs) (S), given in the
CC specification. The novel genes and cells are useful in functional genomic
CC analysis and in the discovery and development of new therapeutic and
CC diagnostic agents and methods. (I) is useful for identifying the coding
CC regions of the murine genome, to isolate cDNAs, genomic clones, or full-
CC length genes/polynucleotides or homologues, heterologues, paralogues, or
CC orthologues that are capable of hybridizing to one or more of the GTSs
CC under stringent conditions. (I) can be incorporated into a phage display
CC system that can be used to screen for proteins, or other ligands, that
CC are capable of binding an amino acid sequence encoded by an
CC oligonucleotide or polynucleotide sequence in at least one of the TS
CC sequences. (I) is useful in addressable arrays, such as gene chips, to
CC identify and characterize temporal and tissue specific gene expression,
CC to identify the gene of interest from many sources and for genetic
CC manipulations such as antisense inhibition and gene targeting. Decreasing
CC the level of expression of (I) and/or down regulating the activity of
CC peptides or proteins encoded by (I) is useful for treating development
CC and cell differentiation disorders, aging, cancer, autoimmune disease,
CC lupus, inflammatory disorders, skin disorders and degenerative disorders.
CC This sequence represents a murine cDNA isolated using gene trap
XX technology
SQ Sequence 407 BP; 82 A; 114 C; 131 G; 74 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 4,926-10 Length: 407
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-6 (1-21) x ABS69736 (1-407)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 283 CAGCAGTGGGAGACGGACCTACCTAGGCAACGCCCTGGTTGTACCTGCTATGGAGGAAGC 342

Qy 21 Arg 21
Db 343 CGG 345

RESULT 2
AAFI8086
ID AAFI8086 standard; DNA; 867 BP.
XX
XX AAFI8086;
AC
XX
DT 14-MAR-2001 (first entry)
DE Lung cancer associated polynucleotide sequence SEQ ID 105.
XX
XX Human; lung cancer associated protein; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active; vulnerary;
KW gastrointestinal; nephrotropic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease; ds.
XX
OS Homo sapiens.
XX WO200055180-A2.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005918.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX Ruben SM;
XX WPI; 2000-587514/55.
XX P-PSDS; AAB58210.
XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.
XX Claim 1; Page 576-577; 1425pp; English.
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective, cytostatic, cardioactive, and
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
SQ Sequence 867 BP; 195 A; 223 C; 268 G; 176 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 1,22e-09 Length: 867
Score: 122.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-581-651D-6 (1-21) x AAF18086 (1-867)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 448 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTTACTTGTATGGAGGAAGC 507

Qy 21 Arg 21
Db 508 CGA 510

RESULT 3
ADSI17488
ID ADSI17488 standard; DNA; 1929 BP.
XX
XX ADSI17488;
AC
XX
DT 02-DEC-2004 (first entry)
XX Nucleotide sequence of human fibronectin 1.
XX cell state; time-lapse profile; protein-protein interaction;
KW

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 562.251 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-6

Perfect score: 122
Sequence: 1 QQMERTYLGALVCTCYGGR 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=GenEmbl -QWRT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1.1 9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	122	100.0	706 6 CQ731570 Sequence
2	122	100.0	1929 6 CQ871810 Sequence
3	122	100.0	1929 6 CQ871828 Sequence
4	122	100.0	2147 6 BD137021 Polypepti

5	122	100.0	2147	6	AX003229	Sequence
6	122	100.0	2147	9	HS276395	Homo sapi
7	122	100.0	2192	9	HS276395	Homo sapi
8	122	100.0	2358	6	CQ875358	Sequence
9	122	100.0	3059	10	BC051082	Mus muscu
10	122	100.0	3796	10	RNF181	Rat fibron
11	122	100.0	4321	9	HS276395	Homo sapi
12	122	100.0	6816	6	CQ715726	Sequence
13	122	100.0	7502	9	HS276395	Homo sapi
14	122	100.0	7544	9	HS276395	Homo sapi
15	122	100.0	7753	9	AB191261	Homo sapi
16	122	100.0	7868	9	HS276395	Homo sapi
17	122	100.0	7951	9	HS276395	Homo sapi
18	122	100.0	7951	9	HS276395	Homo sapi
19	122	100.0	8027	6	CQ833991	Sequence
20	122	100.0	8027	6	CQ875357	Sequence
21	122	100.0	8030	9	HS276395	Homo sapi
22	122	100.0	8035	9	HS276395	Homo sapi
23	122	100.0	8042	9	HS276395	Homo sapi
24	122	100.0	8044	6	AR454662	Sequence
25	122	100.0	8044	6	AX281712	Sequence
26	122	100.0	8320	9	HS276395	Homo sapi
27	122	100.0	8329	6	AX402055	Sequence
28	122	100.0	8329	10	RNF181	Rat mRNA fo
29	122	100.0	8411	9	HS276395	Homo sapi
30	122	100.0	8421	9	HS276395	Homo sapi
31	122	100.0	143947	9	HS276395	Homo sapi
32	122	100.0	185651	2	AC124821	Mus muscu
33	122	100.0	197121	2	AC127101	Rattus no
34	122	100.0	204737	2	AC091456	Mus muscu
35	122	100.0	255929	2	AC095457	Rattus no
36	118	96.7	7679	6	AR274901	Sequence
37	118	96.7	7680	6	AR380744	Sequence
38	118	96.7	7680	6	AX277596	Sequence
39	118	96.7	7680	6	AX335368	Sequence
40	118	96.7	7705	6	HS276395	Human mRNA
41	118	96.7	7705	6	AL14133	Fibronectin
42	118	96.7	7705	6	AR034630	Sequence
43	118	96.7	7705	6	E01162	cdna encodi
44	118	96.7	7705	6	I70110	Sequence 16
45	118	96.7	7705	6	AR364992	Sequence

ALIGNMENTS

RESULT 1	CQ731570	706 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	Sequence 17504 from Patent WO02068579.				
DEFINITION	CQ731570				
ACCESSION	CQ731570.1	GI:42308500			
VERSION					
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.				
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof				
JOURNAL	Patent: WO 02068579-A 17504 06-SEP-2002; PE Corporation (NY) (US)				
FEATURES	Location/Qualifiers				
source	1..706				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Alignment Scores:					
Pred. No.:	1.15e-09				706
Score:	122.00				21

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-581-651D-6 (1-21) x CQ731570 (1-706)

Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 187 CAACAGTGGGAGCGGACCTACCTAGCAATCGTTGGTTTGTACTTGTATGGAGGAGC 246
Qy 21 Arg 21
Db 247 CGA 249

RESULT 2
LOCUS CQ871810 1929 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 1 from Patent WO2004079332.
ACCESSION CQ871810
VERSION CQ871810.1 GI:52745842
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Miyake,M.A., Yoshikawa,T.A., Uchimura,E.A. and Miyake,J.A.
TITLE Composition and method for increasing efficiency of introduction of
JOURNAL target substance into cell
Patent: WO 2004079332-A 1 16-SEP-2004;
National Institute of Advanced Industrial Science and Tec hnology
(JP)
FEATURES
source Location/Qualifiers
1..1929
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1929
/note="unnamed protein product; fibronectin 1"
CDS
/protein_id="CAH56867.1"
/db_xref="GI:52745843"
/translation="MLRPGPGLLLAVQCLGTAVPSTGASKSKRQAQMVQPSVA
VSQKPGCYDNGKHQYINQWERTYLGNALVCTCYGSGRGFNCESKPEAETCFDKYT
GNTVRGDTYERPKDSMTWDCTICAGRGRICTIANRCHGGGOSYKIGDTWRPHET
GGYMLECVCLNGKGEWTKCPAEKCFDHAAGTSYVVGTEWKEPYQGMMVDCITLGE
GSGRTICTSRNCRNDQRTSYRIGDTWSKDNRLQLQICITGNRGWECERHTSV
QTTSSGSGPFTDRAAVYQPPHPPPYGHCVTDSSGVVSVGMQLTKQGNKMLCT
CLNGVSCQETAVTQTYGNSNGEPCVLPTFYNGRTDSTTSNEDQDKYSFCTDHTVL
VOTRGNSGALCHFPFLYNNHNYDCTSEGRDNMKWCGTTONYDADQKFCPMAA
HEEICTTNEGVMYRIGDQWDKQHDHGMWRCCTCVNGRGWETCTAYSQLRQCIYDDI
TYVNDTFHKRHEGHEMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYQIGDSWEKYVH
GVRYQCYCGRGIGGWHCQPLQTYFPSSSGPVEVFTETPSPNSHPIQWNAPOPSHIS
KYLRLWRPVSIPRNLGY"
ORIGIN
Alignment Scores: 2.78e-09 Length: 1929
Pred. No.: 122.00 Matches: 21
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-09-581-651D-6 (1-21) x CQ871828 (1-1929)
Qy 1 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 20
Db 187 CAACAGTGGGAGCGGACCTACCTAGCAATCGTTGGTTTGTACTTGTATGGAGGAGC 246
Qy 21 Arg 21
Db 247 CGA 249

RESULT 4
LOCUS BD137021 2147 bp DNA linear PAT 18-SEP-2002
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION BD137021
VERSION BD137021.1 GI:23231966
KEYWORDS JP 2002508179-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2147)
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 6.14379 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRPVSIPPNLGY 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	749	2 Q7Z391	Q7Z391 homo sapien
2	61	55.5	2265	1 F1NC_BOVIN	P07589 bos taurus
3	56	50.9	1103	2 Q6MZF4	Q6mzf4 homo sapien
4	56	50.9	2193	2 Q6MZM7	Q6mzm7 homo sapien
5	56	50.9	2240	2 Q6MDP8	Q6mdp8 homo sapien
6	56	50.9	2267	2 Q6SDP9	Q6sdp9 homo sapien
7	56	50.9	2296	2 Q6N0A6	Q6n0a6 homo sapien
8	56	50.9	2357	2 Q68D74	Q68dt4 homo sapien
9	56	50.9	2386	1 F1NC_HUMAN	P02751 homo sapien
10	56	50.9	2444	2 Q6N025	Q6n025 homo sapien
11	56	50.9	2477	2 Q6MZU5	Q6mzu5 homo sapien
12	53	48.2	1367	1 IG1R_HUMAN	P08069 homo sapien
13	53	48.2	2477	1 F1NC_MOUSE	P11276 mus musculus
14	53	48.2	2477	1 F1NC_RAT	P04937 rattus norv
15	52	47.3	1358	2 Q73798	Q73798 xenopus lae
16	51	46.4	196	2 Q57542	Q57542 oenococcus
17	51	46.4	1253	2 Q57566	Q57566 canis famil
18	51	46.4	2936	2 Q7YR88	Q7yrk8 canis famil
19	50	45.5	275	2 Q825H3	Q825h3 streptomyce
20	50	45.5	293	2 Q9HTT0	Q9htt0 pseudomonas
21	50	45.5	485	2 Q8H633	Q8h633 oryza sativ
22	50	45.5	1370	1 IG1R_RAT	P24062 rattus norv
23	50	45.5	1371	2 Q9QVW4	Q9qvz4 rattus sp.
24	50	45.5	1373	1 IG1R_MOUSE	Q60751 mus musculu
25	49.5	45.0	372	2 Q9HGK6	Q9hgv6 salicella c
26	48	43.6	189	2 Q8BBK5	Q8bbk5 hepatitis c
27	48	43.6	542	2 Q8Y2T1	Q8y2t1 raietonia s
28	48	43.6	2515	2 Q77365	Q77365 plasmodium s
29	47.5	43.2	922	2 Q93405	Q93405 brachydanio
30	47.5	43.2	2478	2 Q93406	Q93406 brachydanio
31	47	42.7	126	1 PBE3_PHOLL	Q7n3r1 photorhabdu

32 47 42.7 260 2 Q88CN7
33 47 42.7 283 2 Q6H4G2
34 47 42.7 439 1 HGD_DROME
35 47 42.7 645 2 Q8PUN5
36 46.5 42.3 241 2 Q6SHD4
37 46 41.8 87 2 Q8H444
38 46 41.8 196 1 K18A_MOUSE
39 46 41.8 196 2 Q6PDU2
40 46 41.8 239 2 Q8S2P6
41 46 41.8 471 2 Q8P3X6
42 46 41.8 471 2 Q8PFG0
43 46 41.8 539 2 Q84J36
44 46 41.8 610 2 Q8A6W7
45 46 41.8 1407 2 Q6SK65

Q88cn7 pseudomonas
Q6h4g2 oryza sativ
Q9vkj0 drosophila
Q8pun5 methanosarc
Q6shd4 uncultured
Q8h444 oryza sativ
Q9n85 mus musculus
Q6pdu2 rattus norv
Q8s2p6 oryza sativ
Q8p3x6 xanthomonas
Q8pfg0 xanthomonas
Q84j36 oryza sativ
Q8a6w7 bacteroides
Q6sk65 arthrobacte

ALIGNMENTS

RESULT 1
Q7Z391
ID Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecher H., Boscher M., Mewes H.W., Weil B., Amid C., Oeanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn.
DR Pfam; PF00039; fnl; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 110; DB 2; Length 749;

Best Local Similarity 100.0%; Pred.No. 2.2e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLWRPVSIPPNLGY 20

Db 730 ISKYLWRPVSIPPNLGY 749

RESULT 2

F1NC_BOVIN STANDARD; PRT; 2265 AA.

ID F1NC_BOVIN AC P07589;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Fibronectin (FN).

GN Name=FN1;

OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RA MEDLINE=87054047; PubMed=3780752;
 RX Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;
 RL "Complete primary structure of bovine plasma fibronectin.";
 RN Eur. J. Biochem. 161:441-453(1986).
 RP [2]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=83117805; PubMed=6218503;
 RA Petersen T.E., Thorgersen H.C., Skorstengaard K., Vibe-Pedersen K.,
 RA Sahl P., Sottrup-Jensen L., Magnusson S.;
 RL "Partial primary structure of bovine plasma fibronectin: three types
 of internal homology.";
 RN Proc. Natl. Acad. Sci. U.S.A. 80:137-141(1983).
 RP [3]
 RP SEQUENCE OF 2170-2285 FROM N.A.
 RX MEDLINE=83221567; PubMed=6304699;
 RA Kornblith A.R., Vibe-Pedersen K., Baralle F.E.;
 RL "Isolation and characterization of cDNA clones for human and bovine
 fibronectins";
 RN Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222(1983).
 CC -|- FUNCTION: Fibronectins bind cell surfaces and various compounds
 including collagen, fibrin, heparin, DNA, and actin. Fibronectins
 are involved in cell adhesion, cell motility, opsonization, wound
 healing, and maintenance of cell shape.
 CC -|- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
 variants, connected by 2 disulfide bonds near the carboxyl ends;
 to a lesser extent homodimers.
 CC -|- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -|- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=1;
 Comment=A number of isoforms are produced. Each of the "extra
 domain" and the connecting strand 3 are present in some forms of
 fibronectin and absent in others;
 Name=1;
 IsoId=P07589-1; Sequence=Displayed;
 CC -|- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
 by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
 forms), made by fibroblasts, epithelial and other cell types, is
 deposited as fibrils in the extracellular matrix.
 CC -|- PTM: Sulfated (By similarity).
 CC -|- SIMILARITY: Contains 12 fibronectin type I domains.
 CC -|- SIMILARITY: Contains 2 fibronectin type II domains.
 CC -|- SIMILARITY: Contains 15 fibronectin type III domains.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL; K00800; AAA30521.2; -.
 DR PIR; A26452; FNBO.
 DR HSP; P08253; 1KSO.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR000083; Fibrnctnl.
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR008957; FN III-like.
 DR InterPro; IPR000562; FN Type II.
 DR InterPro; IPR003962; FnIII_subd.
 DR Pfam; PF00039; fn1; 12.
 DR Pfam; PF00040; fn2; 2.
 DR Pfam; PF00041; fn3; 15.
 DR PRINTS; PR00012; FNTYPEI.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00014; FNTYPEIII.

DR ProDom; PD000995; FN Type II; 2.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01253; FIBRONECTIN_1; 12.
 DR PROSITE; PS00023; FIBRONECTIN_2; 2.
 DR PROSITE; PS00853; FN3; 15.
 KW Acute phase; Alternative splicing; Cell adhesion; Cell shape;
 KW Direct protein sequencing; Glycoprotein; Heparin-binding;
 KW Phosphorylation; Plasma; Pyrrolidone carboxylic acid; Repeat;
 KW Sulfation.
 FT MOD_RES 1 1 Pyrrolidone carboxylic acid.
 FT DOMAIN 21 241 Fibrin- and heparin-binding 1.
 FT DOMAIN 277 577 Collagen-binding.
 FT DNA_BIND 876 1141 Cell-attachment.
 FT DOMAIN 1236 1509 Heparin-binding 2.
 FT DOMAIN 1600 1870 Fibrin-binding 2.
 FT DOMAIN 1991 2216 Fibronectin type-I 1.
 FT DOMAIN 19 59 Fibronectin type-I 2.
 FT DOMAIN 64 107 Fibronectin type-I 3.
 FT DOMAIN 108 151 Fibronectin type-I 4.
 FT DOMAIN 153 197 Fibronectin type-I 5.
 FT DOMAIN 198 242 Fibronectin type-I 6.
 FT DOMAIN 275 314 Fibronectin type-I 1.
 FT DOMAIN 314 373 Fibronectin type-II 1.
 FT DOMAIN 374 438 Fibronectin type-II 2.
 FT DOMAIN 437 480 Fibronectin type-I 7.
 FT DOMAIN 485 527 Fibronectin type-I 8.
 FT DOMAIN 528 571 Fibronectin type-I 9.
 FT DOMAIN 576 668 Fibronectin type-III 1.
 FT DOMAIN 689 778 Fibronectin type-III 2.
 FT DOMAIN 780 867 Fibronectin type-III 3.
 FT DOMAIN 877 964 Fibronectin type-III 4.
 FT DOMAIN 965 1053 Fibronectin type-III 5.
 FT DOMAIN 1056 1141 Fibronectin type-III 6.
 FT DOMAIN 1142 1234 Fibronectin type-III 7.
 FT DOMAIN 1235 1325 Fibronectin type-III 8.
 FT DOMAIN 1326 1415 Fibronectin type-III 9.
 FT DOMAIN 1416 1505 Fibronectin type-III 10.
 FT DOMAIN 1510 1599 Fibronectin type-III 11 (extra domain).
 FT DOMAIN 1602 1689 Fibronectin type-III 12.
 FT DOMAIN 1692 1780 Fibronectin type-III 13.
 FT DOMAIN 1781 1870 Fibronectin type-III 14.
 FT DOMAIN 1871 1990 Connecting strand 3 (CS-3) (V region).
 FT DOMAIN 1979 2069 Fibronectin type-III 15.
 FT DOMAIN 2083 2127 Fibronectin type-I 10.
 FT DOMAIN 2128 2170 Fibronectin type-I 11.
 FT DOMAIN 2172 2215 Fibronectin type-I 12.
 FT SITE 1493 1495 Cell attachment site.
 FT DISULFID 21 47
 FT DISULFID 45 56
 FT DISULFID 66 94
 FT DISULFID 92 104
 FT DISULFID 110 138
 FT DISULFID 136 148
 FT DISULFID 155 184
 FT DISULFID 182 194
 FT DISULFID 200 229
 FT DISULFID 227 239
 FT DISULFID 277 304
 FT DISULFID 302 311
 FT DISULFID 329 355
 FT DISULFID 343 370
 FT DISULFID 389 415
 FT DISULFID 403 430
 FT DISULFID 439 467
 FT DISULFID 465 477
 FT DISULFID 487 514
 FT DISULFID 512 524
 FT DISULFID 530 558
 FT DISULFID 556 568
 FT DISULFID 2085 2114
 FT DISULFID 2112 2124
 FT DISULFID 2130 2157
 FT DISULFID 2155 2167

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 1.51634 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLKRPVSPRNLGY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	55.5	2265	1 FNBO	fibronectin - bovi
2	56	50.9	2386	1 FNHU	fibronectin precu
3	53	48.2	1367	1 TGHUT1	insulin-like growt
4	53	48.2	2477	2 S14428	fibronectin precu
5	51	46.4	196	2 S42038	hypothetical prote
6	50	45.5	293	2 B92988	hypothetical prote
7	50	45.5	1371	2 A33837	insulin-like growt
8	48	43.6	2523	2 T18477	hypothetical prote
9	45	40.9	203	2 A75453	conserved hypothet
10	45	40.9	224	2 D82966	conserved hypothet
11	44	40.0	196	2 A69193	hypothetical prote
12	44	40.0	210	2 I55609	insulin receptor -
13	44	40.0	719	2 T09543	deoxyxylulose synt
14	44	40.0	958	2 T04870	cellulose synthase
15	44	40.0	1372	2 A34157	insulin receptor p
16	44	40.0	1382	1 INHUR	insulin receptor p
17	44	40.0	1383	2 A36080	insulin receptor p
18	44	40.0	2481	2 A43908	fibronectin - Afri
19	44	40.0	2944	2 A54849	collagen alpha 1(V
20	43.5	39.5	429	2 D82669	citrate synthase X
21	43.5	39.5	442	2 S20908	tubulin beta chain
22	43.5	39.5	447	2 J00423	tubulin beta chain
23	43.5	39.5	447	2 A25377	tubulin beta chain
24	43	39.1	219	2 T46472	hypothetical prote
25	43	39.1	321	2 H70669	hypothetical prote
26	43	39.1	460	2 D96813	hypothetical prote
27	43	39.1	512	1 NTAIWB	nitrogenase (EC 1.
28	43	39.1	512	2 A51986	nitrogenase molybd
29	43	39.1	561	2 B90414	hypothetical prote

RESULT 1

FNBO

fibronectin - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1988 #sequence, revision 31-Dec-1988 #text_change 09-Jul-2004

C:Accession: A26452; B21165; A23292

R:Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.

Eur. J. Biochem. 161, 441-453, 1986

A:Title: Complete primary structure of bovine plasma fibronectin.

A:Reference number: A26452; MUID:87054047; PMID:3780752

A:Accession: A26452

A:Molecule type: protein

A:Residues: 1-2265 <SKO>

A:Cross-references: UNIPROT:P07589

R:Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983

A:Title: Isolation and characterization of cDNA clones for human and bovine fibronectins

A:Reference number: A21165; MUID:83221567; PMID:6304699

A:Accession: B21165

A:Molecule type: mRNA

A:Residues: 2170-2265 <KOR>

A:Cross-references: GB:K00800; NID:G163055; PIDN:AAA30521.2; PID:G5713323

R:Petersen, T.E.; Thogersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sottr

Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983

A:Title: Partial primary structure of bovine plasma fibronectin: three types of internal

A:Reference number: A23292; MUID:83117805; PMID:6218503

A:Accession: A23292

A:Molecule type: protein

A:Residues: 1-16, 'C', 18-20, 'S', 22-432; 447-463; 1367-1517; 1567-1673; 2062-2176, 'N', 2178-226

C:Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.

C:Comment: The plasma fibronectin molecule consists of two chains, which are connected b

C:Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib

aling, and maintenance of cell shape.

C:Comment: Plasma fibronectin is synthesized by hepatocytes.

C:Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe

C:Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu

F:21-241/Domain: fibrin and heparin binding <FBR>

F:21-56/Domain: fibronectin type I repeat homology <1F1>

F:66-104/Domain: fibronectin type I repeat homology <1F2>

F:110-148/Domain: fibronectin type I repeat homology <1F3>

F:155-194/Domain: fibronectin type I repeat homology <1F4>

F:200-239/Domain: fibronectin type I repeat homology <1F5>

F:277-577/Domain: collagen binding <CBR>

F:277-311/Domain: fibronectin type I repeat homology <1F6>

F:329-370/Domain: fibronectin type II repeat homology <2F1>

F:389-430/Domain: fibronectin type II repeat homology <2F2>

F:439-477/Domain: fibronectin type I repeat homology <1F7>

F:487-524/Domain: fibronectin type I repeat homology <1F8>

F:530-568/Domain: fibronectin type I repeat homology <1F9>

F:578-661/Domain: fibronectin type III repeat homology <FN3A>

F:688-770/Domain: fibronectin type III repeat homology <FN3B>

F:779-860/Domain: fibronectin type III repeat homology <FN3C>

sensory transducti
probable ATP-depen
cellulose synthase
C29E4.4 protein -
ORF6 protein - Orf
tubulin beta-2 cha
tubulin beta chain
aminoglycoside 6'-
hypothetical prote
hypothetical prote
hypothetical prote
copalt transport A
hypothetical prote
probable permease
branched-chain ami
hypothetical prote

ALIGNMENTS

F:875-957/Domain: fibronectin type III repeat homology <FN3D>
F:965-1046/Domain: fibronectin type III repeat homology <FN3E>
F:1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F:1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F:1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F:1326-1404/Domain: fibronectin type III repeat homology <FN3I>
F:1410-1517/Domain: cell attachment <CAD>
F:1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F:1493-1495/Region: cell attachment (R-G-D) motif
F:1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F:1600-1870/Domain: heparin binding <HB2>
F:1600-1682/Domain: fibronectin type III repeat homology <FN3L>
F:1692-1773/Domain: fibronectin type III repeat homology <FN3M>
F:1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F:1970-1972/Region: cell attachment (R-G-D) motif
F:1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F:1985-2216/Domain: fibrin binding <FB2>
F:2085-2124/Domain: fibronectin type I repeat homology <1F10>
F:2130-2167/Domain: fibronectin type I repeat homology <1F11>
F:2174-2209/Domain: fibronectin type I repeat homology <1F12>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:3/Cross-link: isopeptide (Gln) (interchain to lys N6-amino of fibrin) #status expirime
F:21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-3
7,2155-2167,2174-2200,2198-2209/Diulfide bonds: #status predicted
F:399,497,511,846,976,1213,1987/Binding site: carbohydrate (Asn) (covalent) #status expe
F:1205,1692/Binding site: carbohydrate (Asn) (covalent) #status absent
F:1943,1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F:2246/bisulfide bonds: interchain (to 2250) #status predicted
F:2250/bisulfide bonds: interchain (to 2246) #status predicted
F:2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 55.5%; Score 61; DB 1; Length 2265;
Best Local Similarity 68.8%; Pred. No. 0.78;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ISKYLWRPVSIPPR 16
|||||||:|:|
DB 607 ISKYLWRKPNSPDR 522

RESULT 2
FNHU
fibrinectin precursor [validated] - human
N/Alternate names: fibronectin splice form ED-A
C/Species: Homo sapiens (man)
C/Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004
C/Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22
R/Dean, D.C.; Bowlus, C.L.; Bourgeois, S.
Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987
A/Title: Cloning and analysis of the promoter region of the human fibronectin gene.
A/Reference number: A26460; MUID:87175578; PMID:3031656
A/Accession: A26460
A/Molecule type: DNA
A/Residues: 1-49 <DEA>
A/Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA5337
R/Oldberg, A.; Ruoslahti, E.
J. Biol. Chem. 261, 2113-2116, 1986
A/Title: Evolution of the fibronectin gene.
A/Reference number: A26284; MUID:86111901; PMID:3003095
A/Accession: A26284
A/Molecule type: DNA
A/Residues: 1447-1540 <OLD>
A/Cross-references: GB:M12549; NID:g182688
A/Note: the authors translated the codon TTC for residue 1494 as Glu
R/Paoletti, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.
Nucleic Acids Res. 16, 3545-3557, 1988
A/Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B
A/Reference number: S00848; MUID:88233940; PMID:3375063
A/Accession: S03917
A/Molecule type: DNA
A/Residues: 1594-1767 'V', 1769-1783 <PAO>
A/Cross-references: EMBL:X07718; NID:g31402
A/Note: the authors translated the codon AAC for residue 1631 as Asp

R/Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.
FEBS Lett. 207, 287-291, 1986
A/Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:
A/Reference number: A24854; MUID:87030929; PMID:3770201
A/Accession: A24854
A/Molecule type: DNA
A/Residues: 1992-2147 <VIB>
A/Cross-references: GB:X04530; NID:g31436
R/Gutman, A.; Yamada, K.M.; Kornblitt, A.
FEBS Lett. 207, 145-148, 1986
A/Title: Human fibronectin is synthesized as a pre-propolypeptide.
A/Reference number: A24476; MUID:87030890; PMID:3770189
A/Accession: A24476
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-14, 'Q', 16-38 <GT>
R/Kornblitt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A/Title: Primary structure of human fibronectin: differential splicing may generate at l
A/Reference number: A91008; MUID:85284965; PMID:2992939
A/Accession: A91008
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 32-1344, 1346-2080; 2112-2386 <KOR>
R/Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A/Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid
A/Reference number: A93529; MUID:84272258; PMID:6462919
A/Accession: A93529
A/Molecule type: mRNA
A/Residues: 973-2080; 2112-2386 <KO2>
A/Cross-references: GB:X00739
R/Oldberg, A.; Limney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A/Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A/Reference number: A21011; MUID:83290929; PMID:6688418
A/Accession: A21011
A/Molecule type: mRNA
A/Residues: 1434-1537 <OL2>
A/Cross-references: GB:K00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R/Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A/Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A/Reference number: A90495; MUID:85280409; PMID:2992573
A/Accession: A90495
A/Molecule type: mRNA
A/Residues: 1594-2386 <BER>
A/Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
R/Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A/Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A/Reference number: A22245; MUID:85231203; PMID:2989004
A/Accession: A22245
A/Molecule type: mRNA
A/Residues: 1948-2067 <UME>
A/Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
A/Accession: B22245
A/Molecule type: mRNA
A/Residues: 1975-1991; 2017-2039 <UM2>
A/Cross-references: GB:M27590
R/Sekiguchi, K.; Klog, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A/Title: Human liver fibronectin complementary DNAs: identification of two different mes
A/Reference number: I52394; MUID:87026578; PMID:3021206
A/Accession: I65273
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1978-1990, 2016-2018, 'N', 2020-2081, 2113-2127 <SEK>
A/Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
R/Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A/Title: Isolation and characterization of cDNA clones for human and bovine fibronectin:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 6.77124 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLWRPVSIPPRNLGY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB.pap.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB.pap.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pap.*
4: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pap.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pap.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pap.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pap.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pap.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pap.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pap.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pap.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pap.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pap.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pap.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pap.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pap.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pap.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB.pap.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pap.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	642	16	US-10-741-601-354
2	110	100.0	642	17	US-10-741-600-1066
3	110	100.0	657	16	US-10-741-601-359
4	110	100.0	657	17	US-10-741-600-1072
5	56	50.9	26	13	US-10-156-820-2
6	56	50.9	26	13	US-10-156-820-78
7	56	50.9	75	14	US-10-005-171-1
8	56	50.9	75	15	US-10-431-642-1
9	56	50.9	984	16	US-10-741-601-356
10	56	50.9	984	17	US-10-741-600-1069
11	56	50.9	1173	18	US-10-450-763-52634

Sequence 4, Appli
Sequence 363, App
Sequence 1075, Ap
Sequence 8, Appli
Sequence 2, Appli
Sequence 64, Appl
Sequence 70, Appl
Sequence 98, Appl
Sequence 98, Appl
Sequence 677, App
Sequence 677, App
Sequence 677, App
Sequence 104, App
Sequence 235, App
Sequence 3, Appli
Sequence 94, Appl
Sequence 357, App
Sequence 366, App
Sequence 104, App
Sequence 1067, App
Sequence 1078, App
Sequence 147, App
Sequence 436, App
Sequence 1137, App
Sequence 28, Appl
Sequence 1, Appli
Sequence 360, App
Sequence 206, App
Sequence 32, Appl
Sequence 1071, App
Sequence 59, Appl
Sequence 4, Appli
Sequence 9, Appli
Sequence 53638, A

56 50.9 2220 15 US-10-236-392-4
56 50.9 2296 16 US-10-741-601-363
56 50.9 2296 17 US-10-741-600-1075
56 50.9 2320 14 US-10-279-733-8
56 50.9 2320 15 US-10-236-392-2
56 50.9 2328 14 US-10-171-311-64
56 50.9 2328 15 US-10-236-031B-70
56 50.9 2328 15 US-10-374-979-98
56 50.9 2328 15 US-10-182-936A-98
56 50.9 2328 16 US-10-477-238A-677
56 50.9 2328 16 US-10-680-287A-677
56 50.9 2328 17 US-10-477-173-677
56 50.9 2355 15 US-10-144-194A-104
56 50.9 2355 15 US-10-360-101-235
56 50.9 2355 15 US-10-447-161-3
56 50.9 2355 16 US-10-734-564-94
56 50.9 2355 16 US-10-741-601-357
56 50.9 2355 16 US-10-741-601-366
56 50.9 2355 16 US-10-491-566-104
56 50.9 2355 17 US-10-741-600-1067
56 50.9 2355 17 US-10-741-600-1078
56 50.9 2355 17 US-10-852-335A-147
56 50.9 2355 18 US-10-287-436A-436
56 50.9 2355 18 US-10-287-436A-1137
56 50.9 2355 20 US-11-040-130-28
56 50.9 2386 10 US-09-561-403-1
56 50.9 2386 16 US-10-741-601-360
56 50.9 2386 16 US-10-788-792-206
56 50.9 2386 16 US-10-618-281-32
56 50.9 2386 17 US-10-741-600-1071
56 50.9 2386 17 US-10-868-577A-59
56 50.9 2386 18 US-10-485-758-4
56 50.9 2386 18 US-10-485-758-9
56 50.9 2474 18 US-10-450-763-52638

ALIGNMENTS

RESULT 1
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 100.0%; Score 110; DB 16; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ISKYLWRPVSIPPRNLGY 20
Db 623 ISKYLWRPVSIPPRNLGY 642

RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match 100.0%; Score 110; DB 17; Length 642;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRRPVSIPPRNLGY 20
|||
Db 623 ISKYLRRPVSIPPRNLGY 642

RESULT 3
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Query Match 100.0%; Score 110; DB 16; Length 657;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRRPVSIPPRNLGY 20
|||
Db 638 ISKYLRRPVSIPPRNLGY 657

RESULT 4
US-10-741-600-1072
; Sequence 1072, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1072
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1072

Query Match 100.0%; Score 110; DB 17; Length 657;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 ISKYLRRPVSIPPRNLGY 20
|||
Db 638 ISKYLRRPVSIPPRNLGY 657

RESULT 5
US-10-156-820-2
; Sequence 2, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-156-820-2

Query Match 50.9%; Score 56; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRRP 10
|||
Db 15 ISKYLRRP 24

RESULT 6
US-10-156-820-78
; Sequence 78, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotome
US-10-156-820-78

Query Match 50.9%; Score 56; DB 13; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLRRP 10
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:22:00 ; Search time 2.11765 Seconds
(without alignments)
705.019 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYILWRVPISPPRNLYG 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	50.9	26	4	US-09-155-613A-2
2	56	50.9	26	4	US-09-155-613A-78
3	56	50.9	31	1	US-07-829-462-2
4	56	50.9	31	1	US-08-340-812-2
5	56	50.9	31	1	US-08-459-064B-2
6	56	50.9	31	2	US-08-460-421A-2
7	56	50.9	31	2	US-08-717-169-4
8	56	50.9	31	4	US-09-228-901A-4
9	56	50.9	31	5	PCT-US93-00909-2
10	56	50.9	56	2	US-08-717-169-6
11	56	50.9	56	4	US-09-228-901A-6
12	56	50.9	66	1	US-08-340-812-12
13	56	50.9	66	1	US-08-459-064B-12
14	56	50.9	66	2	US-08-460-421A-12
15	56	50.9	75	2	US-08-717-169-3
16	56	50.9	75	4	US-09-228-901A-3
17	56	50.9	85	1	US-08-340-812-11
18	56	50.9	85	1	US-08-459-064B-11
19	56	50.9	85	2	US-08-460-421A-11
20	56	50.9	89	1	US-08-340-812-13
21	56	50.9	89	1	US-08-459-064B-13
22	56	50.9	89	2	US-08-460-421A-13
23	56	50.9	89	2	US-07-169-7
24	56	50.9	89	4	US-09-228-901A-7
25	56	50.9	96	2	US-08-717-169-2
26	56	50.9	96	4	US-09-228-901A-2
27	56	50.9	109	1	US-07-829-462-1

Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 16, Appli
Sequence 1, Appli
Sequence 1, Appli
Patent No. 5455158
Patent No. 5455158
Sequence 12, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 26, Appli
Sequence 26, Appli
Sequence 26, Appli

28 56 50.9 109 1 US-08-340-812-1
29 56 50.9 109 1 US-08-459-064B-1
30 56 50.9 109 2 US-08-460-421A-1
31 56 50.9 109 2 US-08-717-169-1
32 56 50.9 109 4 US-09-228-901A-1
33 56 50.9 109 5 PCT-US93-00909-1
34 56 50.9 2231 1 US-08-153-799-16
35 56 50.9 2324 1 US-08-283-857-1
36 56 50.9 2324 5 PCT-US95-09819-1
37 56 50.9 2327 6 5455158-1
38 56 50.9 2327 6 5455158-1
39 56 50.9 2386 2 US-09-016-366A-12
40 56 50.9 2386 4 US-09-361-403-1
41 56 50.9 2446 2 US-08-551-356-2
42 56 50.9 2446 5 PCT-US93-12687-2
43 53 48.2 99 1 US-08-036-210-26
44 53 48.2 99 2 US-08-449-609-26
45 53 48.2 99 4 US-09-361-096A-26

ALIGNMENTS

RESULT 1

US-09-155-613A-2
; Sequence 2, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCES: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155.613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-155-613A-2

Query Match 50.9%; Score 56; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYILWRP 10
|||
Db 15 ISKYILWRP 24

RESULT 2

US-09-155-613A-78
; Sequence 78, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCES: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155.613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005

```
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotope
US-09-155-613A-78

Query Match          50.9%; Score 56; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ISKYILRWRP 10
      |||||
Db      15 ISKYILRWRP 24

RESULT 3
US-07-829-462-2
; Sequence 2, Application US/07829462
; Patent No. 5453489
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki I.
; APPLICANT: Morla, Alex
; TITLE OF INVENTION: NOVEL FRAGMENTS OF FIBRONECTIN AND
; TITLE OF INVENTION: METHODS TO MODULATE THE ROLE OF FIBRONECTIN IN
; TITLE OF INVENTION: EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920131
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,547
; REFERENCE/DOCKET NUMBER: P-LA 9179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-829-462-2

Query Match          50.9%; Score 56; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ISKYILRWRP 10
      |||||
Db      8 ISKYILRWRP 17

RESULT 4
US-07-829-462-2

Query Match          50.9%; Score 56; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ISKYILRWRP 10
      |||||
Db      8 ISKYILRWRP 17
```

```
US-08-340-812-2
; Sequence 2, Application US/08340812
; Patent No. 5629291
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS OF
; TITLE OF INVENTION: MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,812
; FILING DATE: 17-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/021,626
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BELLAS, CHRISTINE M.
; REGISTRATION NUMBER: 34,122
; REFERENCE/DOCKET NUMBER: P-LA 9442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-340-812-2

Query Match          50.9%; Score 56; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ISKYILRWRP 10
      |||||
Db      8 ISKYILRWRP 17

RESULT 5
US-08-459-064B-2
; Sequence 2, Application US/08459064B
; Patent No. 5747452
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: A METHOD OF MODULATING TUMOR CELL MIGRATION
; TITLE OF INVENTION: USING FIBRONECTIN TYPE III PEPTIDES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 7.39869 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLWRVPSIPPNLGY 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	20	2	AAI28902 MSP 1-1alp
2	110	100.0	642	8	ADQ39403 Human myo
3	110	100.0	642	8	ADR67316 Human bla
4	110	100.0	642	8	ADSL17489 Amino aci
5	110	100.0	642	8	ADR97658 Human fib
6	110	100.0	657	8	ADQ39409 Human myo
7	110	100.0	660	2	AAI28901 Human mig
8	61	55.5	2265	8	ADSL17498 Amino aci
9	61	55.5	2265	8	ADR97667 Bovine fi
10	56	50.9	20	2	AAI28911 Peptide e
11	56	50.9	26	2	AAW75324 Human fib
12	56	50.9	26	2	AAW75260 Human fib
13	56	50.9	26	2	AAW70995 Motif 1 o
14	56	50.9	31	2	AAW19970 Fibronect
15	56	50.9	31	2	AAW57192 Fibronect
16	56	50.9	31	2	AAW82986 Human fib
17	56	50.9	56	2	AAW57194 Fibronect
18	56	50.9	66	2	AAW19973 Fibronect
19	56	50.9	66	2	AAW82984 Human fib
20	56	50.9	75	2	AAW57191 Fibronect
21	56	50.9	75	6	ABU63426 Human fib
22	56	50.9	75	8	ADI58820 Angiogene
23	56	50.9	85	2	AAW19972 Fibronect
24	56	50.9	85	2	AAW82983 Human fib
25	56	50.9	89	2	AAW19974 Fibronect

26	56	50.9	89	2	AAW57195	AAW57195 Fibronect
27	56	50.9	89	2	AAW82985	AAW82985 Human fib
28	56	50.9	96	2	AAW57190	AAW57190 Fibronect
29	56	50.9	109	2	AAW39729	AAW39729 Fibret typ
30	56	50.9	109	2	AAW57189	AAW57189 Fibronect
31	56	50.9	109	2	AAW82982	AAW82982 Human fib
32	56	50.9	720	2	AAI28914	AAI28914 Fibronect
33	56	50.9	984	8	ADQ39406	ADQ39406 Human myo
34	56	50.9	993	2	AAW80839	AAW80839 Cathepsin
35	56	50.9	1173	4	ABG22275	ABG22275 Novel hum
36	56	50.9	1179	8	ADP75952	ADP75952 Human min
37	56	50.9	1189	2	AAW80839	AAW80839 Human Ser
38	56	50.9	1223	8	ADP30365	ADP30365 Human sec
39	56	50.9	1359	8	ADP75957	ADP75957 Human leu
40	56	50.9	1381	2	AAW80832	AAW80832 Human Ser
41	56	50.9	2182	8	ADR66462	ADR66462 Human pro
42	56	50.9	2182	8	ADR66120	ADR66120 Human pro
43	56	50.9	2220	6	ABO01289	ABO01289 Human pro
44	56	50.9	2220	8	ADN95950	ADN95950 Human NOV
45	56	50.9	2265	4	AAW38647	AAW38647 Human pol

ALIGNMENTS

RESULT 1

AAI28902
ID AAI28902 standard; peptide; 20 AA.

XX AAI28902;

XX 21-SEP-1999 (first entry)

XX MSF 1-alpha peptide epitope.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;

KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.

XX Synthetic.

OS Homo sapiens.

XX W09931233-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

XX 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX Claim 18; Page 57; 86pp; English.

CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. Sequences AAI28902-906 represent peptide epitopes of MSF against which monoclonal antibodies that are specific to MSF and do not cross-react with fibronectin are raised

XX Sequence 20 AA;

Query Match 100.0%; Score 110; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLWRPVSIPPRNLGY 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 ISKYLWRPVSIPPRNLGY 20
| | | | | | | | | | | | | | | | | | | | | |

RESULT 2
ADQ39403
ID ADQ39403 standard; protein; 642 AA.
XX
XX AC ADQ39403;
XX
XX DT 18-NOV-2004 (first entry)
XX DE Human myocardial infarction-associated gene derived protein, SEQ ID 1066.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human.
XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX
XX Cargill M, Devlin JJ, Iakoubova O;
XX PI WPI; 2004-533949/51.
XX DR N-PSDB; ADQ38575.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 10; SEQ ID NO 1066; 145pp; English.

The invention relates to a novel method for identifying an individual who has an altered risk for developing myocardial infarction. The method comprises detecting a single nucleotide polymorphism (SNP) in any one of the nucleotide sequences given in the specification in the individual's nucleic acids, where the presence of the SNP is correlated with an altered risk for myocardial infarction in the individual. The invention further comprises: an isolated nucleic acid molecule comprising at least 8 contiguous nucleotides where one of the nucleotides is an SNP given in the specification or its complement and encoding any one of the amino acid sequences given in the specification; an isolated polypeptide comprising an amino acid sequence given in the specification; an antibody that specifically binds to the polypeptide or its antigen-binding fragment; an amplified polynucleotide containing an SNP given in the specification and which is between about 16 and 1000 nucleotides in length; a kit for detecting an SNP in a nucleic acid, comprising the polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a nucleic acid molecule; a method of detecting a variant polypeptide; and a method for identifying an agent useful in treating or preventing myocardial infarction. The novel detection method has cardiant activity. The nucleic acids of the invention may be used in gene therapy. The method is useful in identifying an individual who has an increased or decreased risk for developing myocardial infarction and for preparing a composition for treating or preventing myocardial infarction. This sequence represents the protein of a human myocardial infarction-associated gene containing one or more SNPs of the invention. Note: This sequence was not shown in the specification. The sequence has come from an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 642 AA;
Query Match 100.0%; Score 110; DB 8; Length 642;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLWRPVSIPPRNLGY 20
| | | | | | | | | | | | | | | | | | | | | |
Db 623 ISKYLWRPVSIPPRNLGY 642
| | | | | | | | | | | | | | | | | | | | | |

RESULT 3
ADR67316
ID ADR67316 standard; protein; 642 AA.
XX
XX AC ADR67316;
XX DT 02-DEC-2004 (first entry)
XX DE Human bladder cancer associated amino acid sequence.
XX KW bladder cancer tissue; bladder cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO2004076613-A2.
XX PD 10-SEP-2004.
XX PF 24-FEB-2004; 2004WO-DE000364.
XX PR 26-FEB-2003; 2003DE-01009729.
XX
XX (HERR/) HERR A.
XX PA (HINZ/) HINZMANN B.
XX PA (DAHL/) DAHL E.
XX PA (STAU/) STAU E.
XX PA (PILA/) PILARSKY C.
XX PA (SPEC/) SPECHT T.
XX
XX Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;
XX WPI; 2004-653385/63.
XX
XX New nucleic acids, and encoded proteins, from bladder cancer tissue,
XX useful for diagnosis, treatment and in screening for specific binding
XX agents.
XX
XX Claim 2; Fig 3; 112pp; German.

The present invention describes nucleic acids (I) associated with bladder cancer tissue. Also described: (1) peptides and proteins (II) containing an amino acid sequence encoded by (I); (2) a method for diagnosing bladder cancer (BC), or monitoring its progression, that uses (I), (II) or agents (Z) that inhibit, or bind to, (I) or (II), labelled with a reporter; and (3) a method for treating BC that uses (I), (II) or (Z). (I) and (II) have cytostatic activity. (I) and (II) can be used to detect and monitor progression of bladder cancer (BC), or the risk of developing it; to screen for specific binding agents (Z), and to treat BC. (Z) are also useful as diagnostic and therapeutic agents. The present sequence represents a human amino acid sequence associated with bladder cancer, which is used in the exemplification of the present invention.

XX SQ Sequence 642 AA;
Query Match 100.0%; Score 110; DB 8; Length 642;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ISKYLWRPVSIPPRNLGY 20
| | | | | | | | | | | | | | | | | | | | | |
Db 623 ISKYLWRPVSIPPRNLGY 642
| | | | | | | | | | | | | | | | | | | | | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 415.948 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRPVSIPPNLGY 20

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US09581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFWT=fa+cap -SURFIX=p2n.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MAIRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651.cGCG 1.1 8180 @runat_07112005_092223_28789 -NCPU=3 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	110	100.0	362	1	AT263888
C 2	110	100.0	391	1	AA234360
C 3	110	100.0	440	1	AI754331
C 4	110	100.0	670	5	BU620723
C 5	110	100.0	706	4	BM674187
C 6	110	100.0	747	6	CA423317
C 7	110	100.0	967	5	EX402381
C 8	110	100.0	1044	5	BX398837
C 9	110	100.0	1052	5	BX380582

10	65	59.1	658	2	BB625874	BB625874
11	61	55.5	532	5	BQ830132	BQ830132
12	61	55.5	546	5	BQ830316	LL6ln2173
13	61	55.5	562	2	BB709173	BB709173
14	61	55.5	702	9	CE165323	CE165323
15	60	54.5	348	2	BB812986	BB812986
16	60	54.5	397	1	AJ665984	AJ665984
17	60	54.5	607	7	CK450361	CK450361
18	60	54.5	613	9	CL586124	CL586124
19	60	54.5	755	7	CK451343	CK451343
20	60	54.5	841	4	BI684770	BI684770
21	60	54.5	896	7	CN163427	CN163427
22	60	54.5	897	7	CN162046	CN162046
23	59	53.6	192	2	BB361047	BB361047
24	59	53.6	1128	8	CC193783	CC193783
25	58	52.7	226	2	BB394322	BB394322
26	58	52.7	1133	4	BM478486	BM478486
27	58	52.7	1521	7	CK411745	CK411745
28	57.5	52.3	423	2	BB812942	BB812942
29	57.5	52.3	828	1	AI287222	AI287222
30	57	51.8	507	8	BZ695434	BZ695434
31	57	51.8	890	2	BE789442	BE789442
32	56.5	51.4	707	1	AU140730	AU140730
33	56	50.9	210	4	BI040656	BI040656
34	56	50.9	219	4	BI054371	BI054371
35	56	50.9	220	4	BI040247	BI040247
36	56	50.9	260	2	BF880844	BF880844
37	56	50.9	268	2	BF367054	BF367054
38	56	50.9	269	2	BF369623	BF369623
39	56	50.9	295	2	BE812941	BE812941
40	56	50.9	301	1	AA853208	AA853208
41	56	50.9	316	2	BF876014	BF876014
42	56	50.9	320	1	AA302280	AA302280
43	56	50.9	320	5	BQ319620	BQ319620
44	56	50.9	328	2	BF869573	BF869573
45	56	50.9	334	2	BF808854	BF808854

ALIGNMENTS

RESULT 1
AI263888/c

LOCUS q108d11.x1 Soares_NHMPu_S1 Homo sapiens cdna clone IMAGE:1855893
DEFINITION 3', similar to gb:X02761_cdsl FIBRONECTIN PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AI263888.1 GI:3872091
VERSION AI263888

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 362)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL Tumor Gene Index

COMMENT Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco.

FEATURES Location/Qualifiers

1..362
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:1855893"
/tissue_types="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_hosts="DH10B"
/clone_lib="Soares_NHMPu_S1"

/note="Organ: mixed (see below); Vector: pTTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN

Alignment Scores: 2.23e-06 Length: 362
 Pred. NO.: 110.00 Matches: 20
 Score: 110.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-581-651d-5 (1-20) x A1263888 (1-362)

Qy 1 IleSerLysTyrIleLeuArgTrrpArgProValSerIleProProArgAsnLeuGlyTyr 20
 |||||
 Db 214 ATTTCAGTACATTCCTCAGGTGGAGACTGTGAGTATCCACCCAGAAACCTTGGATAC 155

RESULT 2

AA234360/c
 LOCUS AA234360 391 bp mRNA linear EST 06-AUG-1997
 DEFINITION zr72c02.s1 Soares NHMPu S1 Homo sapiens cDNA clone IMAGE:668930 3', similar to gb:X02761_cds1 FIBROBLAST PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AA234360
 VERSION AA234360.1 GI:1858635
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS

1 (bases 1 to 391)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheinberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Watson, K. and Wilson, R.

Washington University School of Medicine
 Contact: Wilson RK

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 678 Std Error: 0.00

Seq primer: -41m3 fwd. ET from Amersham

High quality sequence stop: 376.

Location/Qualifiers

FEATURES

source

1..391
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:5562897"
 /db_xref="taxon:9606"
 /clone="IMAGE:668930"
 /tissue type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab host="DH108"
 /clone lib="Soares NHMPu S1"

/note="Organ: mixed (see below); Vector: pTTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus

NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN

Alignment Scores: 2.44e-06 Length: 391
 Pred. NO.: 110.00 Matches: 20
 Score: 110.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0

US-09-581-651d-5 (1-20) x AA234360 (1-391)

Qy 1 IleSerLysTyrIleLeuArgTrrpArgProValSerIleProProArgAsnLeuGlyTyr 20
 |||||
 Db 205 ATTTCAGTACATTCCTCAGGTGGAGACTGTGAGTATCCACCCAGAAACCTTGGATAC 146

RESULT 3

AI754331/c
 LOCUS AI754331 440 bp mRNA linear EST 20-JUN-2002
 DEFINITION cr23e08.x1 Human bone marrow stromal cells Homo sapiens cDNA clone HMSC cr23e08 3', mRNA sequence.

ACCESSION AI754331
 VERSION AI754331.1 GI:5132595
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 440)

Jia, L., Young, M.F., Powell, J., Yang, L., Ho, N.C., Hotchkiss, R., Robey, P.G. and Francomano, C.A.

Gene expression profile of human bone marrow stromal cells:

high-throughput expressed sequence tag sequencing analysis

Genomics 79 (1), 7-17 (2002)

21686149

11827452

Contact: Libin Jia

Medical Genetics Branch

National Human Genome Research Institute

10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA

Tel: 301-402-4877

Fax: 301-496-7157

Email: libin@helix.nih.gov

DNA Sequencing and analyses by National Institutes of Health

Intramural Sequencing Center (NISC).

Plate: 23 row: e column: 08

Seq primer: -21M13 forward primer (ABI).

FEATURES

source

1..440
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HMSC_cr23e08"
 /sex="mixed"
 /tissue type="bone marrow stroma"
 /dev stage="mixed"
 /lab host="XLI-Blue MRF"/SOIR"

/note="Vector: pBluescript; Site 1: EcoRI; Site 2: XhoI; mRNA made from human bone marrow stroma, cDNA made by oligo-dT priming. Directionally cloned. Size-selected for average insert size >0.5 kb. Library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR). Library supplied by Dr. Libin Jia (NHGRI)"

ORIGIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 102.693 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRVSPRNLGY 20

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=Published Applications NA -QFMT=fastp -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09581651 @CGN 1 1 1041 @runat_07112005_092226_28867
-NCPU=6 -ICPU=1 -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	471	18	US-10-242-535A-42467
2	110	100.0	471	19	US-10-085-783A-42467
3	110	100.0	600	22	US-10-956-157-9523
4	110	100.0	2127	17	US-10-210-120-49
5	110	100.0	2127	22	US-10-956-157-4288
6	110	100.0	2127	24	US-10-909-035-49
7	110	100.0	2443	20	US-10-741-601-70
8	110	100.0	2443	22	US-10-741-600-238
9	110	100.0	2488	20	US-10-741-601-75
10	110	100.0	2488	22	US-10-741-600-244
11	110	100.0	4295	18	US-10-144-194A-51
12	110	100.0	4295	21	US-10-491-566-51
13	110	100.0	87467	20	US-10-741-601-5634
14	110	100.0	87467	22	US-10-741-600-17824
15	59	53.6	937	18	US-10-369-493-27562
c 16	56	50.9	540	22	US-10-695-140-1
c 17	56	50.9	567	20	US-10-767-701-30779
c 18	56	50.9	1892	20	US-10-384-339C-133
c 19	56	50.9	2055	24	US-10-974-558-1
c 20	56	50.9	3522	24	US-10-450-763-22266
c 21	56	50.9	6189	16	US-10-259-273-41
c 22	56	50.9	6189	26	US-11-006-313-41
23	56	50.9	6510	23	US-10-741-601-72
24	56	50.9	6510	22	US-10-741-600-241
25	56	50.9	6988	19	US-10-236-392-1
26	56	50.9	7361	19	US-10-236-392-3
27	56	50.9	7677	22	US-10-956-157-4995
28	56	50.9	7679	22	US-10-831-704-38
29	56	50.9	7680	9	US-09-964-824A-574
30	56	50.9	7680	15	US-10-171-311-63
31	56	50.9	7680	18	US-10-236-031B-69
32	56	50.9	7680	18	US-10-374-979-75
33	56	50.9	7680	19	US-10-182-936A-75
34	56	50.9	7680	19	US-10-641-643-1289
35	56	50.9	7680	20	US-10-717-597-222
36	56	50.9	7680	20	US-10-788-792-79
37	56	50.9	7680	21	US-10-477-238A-654
38	56	50.9	7680	21	US-10-580-287A-654
39	56	50.9	7680	22	US-10-278-698-88
40	56	50.9	7680	22	US-10-278-698-603
41	56	50.9	7680	22	US-10-843-641A-5877
42	56	50.9	7680	22	US-10-477-173-654
43	56	50.9	7680	22	US-10-852-135A-52
44	56	50.9	7705	18	US-10-447-161-4
45	56	50.9	7795	16	US-10-084-817-2

ALIGNMENTS

RESULT 1
US-10-242-535A-42467
; Sequence 42467, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Lew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242.535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42467
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-42467

Alignment Scores:
Pred. No.: 5,23e-08 Length: 471
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-581-651D-5 (1-20) x US-10-242-535A-42467 (1-471)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 363 ATTTCGAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 422

RESULT 2

US-10-085-783A-42467
; Sequence 42467, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42467

; LENGTH: 471
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (7)..(7)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-42467

Alignment Scores:
Pred. No.: 5,23e-08 Length: 471
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 19 Gaps: 0

US-09-581-651D-5 (1-20) x US-10-085-783A-42467 (1-471)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 363 ATTTCGAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 422

RESULT 3

US-10-956-157-9523
; Sequence 9523, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:

; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9523
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-9523

Alignment Scores:
Pred. No.: 6.8e-08 Length: 600
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0

US-09-581-651D-5 (1-20) x US-10-956-157-9523 (1-600)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 396 ATTTCGAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 455

RESULT 4

US-10-210-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:

; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-49

Alignment Scores:
Pred. No.: 2.69e-07 Length: 2127
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-581-651D-5 (1-20) x US-10-210-120-49 (1-2127)

Qy 1 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 1923 ATTTCGAAGTACATTCTCAGGTGGAGACCTGTGAGTATCCACCAGAAACCTTGGATAC 1982

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 20.1307 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKIILWRVSIIPRNGLY 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/h/US09581651/runat_07112005_092224_28801/app_query.fasta_1.1834
-DB=issued_Patents_NA -QFIX=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651@cgn 1.187 @runat_07112005_092224_28801 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	57	51.8	374159	4	US-09-949-016-15868
2	56	50.9	186	1	Sequence 15868, A
3	56	50.9	7679	4	US-08-153-799-5
4	56	50.9	7680	4	Sequence 5, Appli
5	56	50.9	7680	5	Sequence 38, Appl
6	56	50.9	7705	2	Sequence 1289, Ap
7	56	50.9	7705	1	Sequence 6, Appli
8	56	50.9	7705	6	Sequence 16, Appl
9	56	50.9	7705	6	Sequence 16, Appl
10	56	50.9	7803	2	Patent No. 5455158
11	56	50.9	7803	5	Patent No. 5455158-2
12	56	50.9	7989	4	Sequence 1, Appli
					Sequence 10, Appl

C 13	56	50.9	8001	4	US-09-539-601-7	Sequence 7, Appli
C 14	56	50.9	8001	4	US-09-539-601-16	Sequence 16, Appl
C 15	56	50.9	8001	4	US-09-539-601-22	Sequence 22, Appl
C 16	56	50.9	8001	4	US-09-539-601-28	Sequence 28, Appl
C 17	56	50.9	8044	4	US-09-566-921-135	Sequence 135, App
C 18	56	50.9	8637	4	US-09-539-601-4	Sequence 4, Appli
C 19	56	50.9	8638	4	US-10-029-907-6	Sequence 6, Appli
C 20	56	50.9	8638	4	US-10-029-907-7	Sequence 7, Appli
C 21	56	50.9	8638	4	US-10-029-907-24	Sequence 24, Appl
C 22	56	50.9	8638	4	US-10-029-907-25	Sequence 25, Appl
C 23	56	50.9	8639	4	US-10-029-907-1	Sequence 1, Appli
C 24	56	50.9	8642	4	US-10-029-907-2	Sequence 2, Appli
C 25	56	50.9	8643	4	US-10-029-907-4	Sequence 4, Appli
C 26	56	50.9	8648	4	US-10-029-907-5	Sequence 5, Appli
C 27	56	50.9	8649	4	US-09-539-601-13	Sequence 13, Appl
C 28	56	50.9	11076	4	US-09-539-601-1	Sequence 1, Appli
C 29	56	50.9	11076	4	US-09-539-601-19	Sequence 19, Appl
C 30	56	50.9	11076	4	US-09-539-601-25	Sequence 25, Appl
C 31	56	50.9	11076	4	US-09-539-601-31	Sequence 31, Appl
C 32	54	49.1	7475	2	US-08-971-036-1	Sequence 1, Appli
C 33	54	49.1	7475	3	US-09-096-570-1	Sequence 1, Appli
C 34	54	49.1	7475	3	US-09-265-617B-1	Sequence 1, Appli
C 35	53.5	48.6	5137	4	US-09-949-016-1011	Sequence 1011, Ap
C 36	53.5	48.6	5140	4	US-09-949-016-3107	Sequence 3107, Ap
C 37	53.5	48.6	9681	4	US-09-949-016-12753	Sequence 12753, A
C 38	53.5	48.6	9681	4	US-09-949-016-14849	Sequence 14849, A
C 39	53	48.2	601	4	US-09-949-001-203	Sequence 203, App
C 40	53	48.2	601	4	US-09-949-001-477	Sequence 477, App
C 41	53	48.2	2320	3	US-09-202-904A-10	Sequence 10, Appl
C 42	53	48.2	4975	2	US-08-249-687C-1	Sequence 1, Appli
C 43	53	48.2	4989	2	US-08-666-392A-3	Sequence 3, Appli
C 44	53	48.2	4989	2	US-08-625-819-1	Sequence 1, Appli
C 45	53	48.2	4989	3	US-08-755-558-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-15868/c
; Sequence 15868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15868
; LENGTH: 374159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15868

Alignment Scores:	1.46e+04	Length:	374159
Pred. No.:	57.00	Matches:	11
Score:	73.33%	Conservative:	0
Percent Similarity:	73.33%	Mismatches:	4
Best Local Similarity:	73.33%	Indels:	0
Query Match:	51.82%	Gaps:	0
DB:	4		
US-09-581-651D-5 (1-20) x US-09-949-016-15868 (1-374159)			
Qy	2 SerLysTyrIleLeuArgTrpArgProValSerIleProProArg 16		

```

Db 300484 TCAGGATGCTCAGCTGGAGGCGCTGGCTCCATCCACCCAGG 300440
|||||
RESULT 2
US-08-153-799-5
; Sequence 5, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 186 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..40
; OTHER INFORMATION: /function= "oligonucleotide 1"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 41..89
; OTHER INFORMATION: /function= "oligonucleotide 2"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 90..138
; OTHER INFORMATION: /function= "oligonucleotide 3"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 139..186
; OTHER INFORMATION: /function= "oligonucleotide 4"

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..186
; OTHER INFORMATION: /function= "LINKER 5"
US-08-153-799-5
Alignment Scores:
Pred. No.: 1.86 Length: 186
Score: 56.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.91% Indels: 0
DB: 1 Gaps: 0
US-09-581-651D-5 (1-20) x US-08-153-799-5 (1-186)
Qy 1 lIeSerLysTyrlleLeuArgTtpArgPro 10
| | | | | | | | | | | | | | | | | | | | | |
Db 85 ATTCCAAGTACATTCTCAGGTGGAGACCT 114
RESULT 3
US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shivjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38
Alignment Scores:
Pred. No.: 179 Length: 7679
Score: 56.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.91% Indels: 0
DB: 4 Gaps: 0
US-09-581-651D-5 (1-20) x US-09-220-132-38 (1-7679)
Qy 1 lIeSerLysTyrlleLeuArgTtpArgPro 10
| | | | | | | | | | | | | | | | | | | | | |
Db 1832 ATTCCAAGTACATTCTCAGGTGGAGACCT 1861
RESULT 4
US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:30:56 ; Search time 65.5686 Seconds
(without alignments)
1805.663 Million cell updates/sec

Title: US-09-581-651D-5
Perfect score: 110
Sequence: 1 ISKYLWRPVSIPPRNLGY 20

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO.spool_b/US09581651/runat_07112005_092222_28772/app_query.fasta_1.1834
-DB=N_Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651.qcgn 1.1 1052 @runat_07112005_092222_28772 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUB -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 16Dec04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001as:.*
5: Geneseq2001bs:.*
6: Geneseq2002as:.*
7: Geneseq2002bs:.*
8: Geneseq2003as:.*
9: Geneseq2003bs:.*
10: Geneseq2003cs:.*
11: Geneseq2003ds:.*
12: Geneseq2004as:.*
13: Geneseq2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	1929	13 ADS17488	Adsl7488 Nucleotid
2	110	100.0	1929	13 ADS97657	Adr97657 Human fib
3	110	100.0	2127	10 ADD18477	Add18477 Human pro
4	110	100.0	2147	2 AAX81299	Aax81299 Human mig
5	110	100.0	2358	13 ADR67201	Adr67201 Human bla

6	110	100.0	2443	13 ADQ38575	Adq38575 Human SNP
7	110	100.0	2488	13 ADQ38581	Adq38581 Human SNP
8	110	100.0	4295	8 ACC72037	Acc72037 BCU0770 g
9	110	100.0	11026	4 AAK72958	Aak72958 Human lmm
10	110	100.0	31749	4 AAK72959	Aak72959 Human lmm
11	110	100.0	78925	3 AAC89888	Aac89888 Human FN
12	59	53.6	937	13 ADS49132	Ads49132 Bacterial
13	56	50.9	225	10 ADD93721	Add93721 Hepatitis
14	56	50.9	499	10 ACD97182	Acd97182 Human col
15	56	50.9	540	12 ADN96854	Adn96854 Hepatitis
16	56	50.9	1892	6 ABV78249	Abv78249 Hepatitis
17	56	50.9	1892	6 ABZ35825	Abz35825 Hepatitis
18	56	50.9	1892	6 ABX10068	Abx10068 HCV NS3 D
19	56	50.9	1892	6 ABL91790	AbL91790 HCV polyn
20	56	50.9	3522	5 AAS86462	Aas86462 DNA encod
21	56	50.9	3540	12 ADP75954	Adp75954 Human min
22	56	50.9	3669	12 ADP29134	Adp29134 Human sec
23	56	50.9	4080	12 ADP75958	Adp75958 Human leu
24	56	50.9	4860	3 AAA35009	Aaa35009 Human ade
25	56	50.9	6189	10 ADD67965	Add67965 Hepatitis
26	56	50.9	6510	13 ADQ38578	Adq38578 Human SNP
27	56	50.9	6816	10 ABX74443	Abx74443 Human CDN
28	56	50.9	6988	9 ACD06169	Acd06169 Human CDN
29	56	50.9	7020	13 ADR90518	Adr90518 Human NOV
30	56	50.9	7049	12 ADN95947	Adn95947 Human NOV
31	56	50.9	7141	6 AAD25333	Aad25333 Hepatitis
32	56	50.9	7242	13 ADR66637	Adr66637 Human pro
33	56	50.9	7242	13 ADR65953	Adr65953 Human pro
34	56	50.9	7361	9 ACD06170	Acd06170 Human CDN
35	56	50.9	7361	12 ADN95949	Adn95949 Human NOV
36	56	50.9	7550	8 ACC00412	Acc00412 Human cel
37	56	50.9	7677	13 ACN37820	Acn37820 Tumour-as
38	56	50.9	7679	10 ADB31322	Adb31322 Testoster
39	56	50.9	7680	2 AAT17551	Aat17551 Human fib
40	56	50.9	7680	3 AAF21131	Aaf21131 Human low
41	56	50.9	7680	5 ABA82689	Aba82689 Fibronect
42	56	50.9	7680	6 ABL67540	AbL67540 Thyroid c
43	56	50.9	7680	6 ABT11082	Abt11082 Human bre
44	56	50.9	7680	8 ACF03878	Acf03878 Human fib
45	56	50.9	7680	8 ABX10391	Abx10391 DNA encod

ALIGNMENTS

RESULT 1
ADS17488
ID ADS17488 standard; DNA; 1929 BP.
XX
AC ADS17488;
XX
DT 02-DEC-2004 (first entry)
XX
DE Nucleotide sequence of human fibronectin 1.
XX
KW cell state; time-lapse profile; protein-protein interaction;
KW drug screening; cancer; infectious disease; allergy; hypertension;
KW hyperlipaemia; diabetes; cardiac disease; cerebral infarction; dementia;
KW obesity; arteriosclerosis; infertility; mental disease; nervous disease;
KW cataract; progeria; hypersensitivity; ultraviolet radiation; human;
KW fibronectin 1; actin acting substance; transfection array; gene; db.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1929
FT /tag= a
FT /product= "fibronectin 1"
XX
XX WO2004079007-A2.
XX
XX PD 16-SEP-2004.
XX
XX 03-MAR-2004; 2004WO-JP002694.

```
XX 04-MAR-2003; 2003JP-00057870.
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX Miyake M, Yoshikawa T, Uchimura E, Miyake J;
XX WPI; 2004-662438/64.
XX P-PSDB; ADS17489.
XX Presenting a state of a cell, useful for diagnosing and treating a
XX disease, e.g. cancer, infectious disease, allergy, diabetes, dementia,
XX obesity, infertility, or cataract, comprises obtaining a time-lapse
XX profile of the cell.
XX Disclosure; SEQ ID NO 1; 532pp; English.
XX The specification describes a method and system for accurately presenting
XX a state of a cell. The method comprises obtaining a time-lapse profile of
XX the cell by time-lapse monitoring of a gene state associated with at
XX least one gene derived from the cell, and presenting the time-lapse
XX profile. The gene comprises a transcription control sequence, and the
XX gene state includes expression of the gene. The method and system are
XX useful presenting a state of a cell. The method can allow the elucidation
XX of key protein-protein interactions suitable for targeting by drug
XX screening protocols. The method is useful for diagnosing or treating a
XX disease, e.g. cancer, infectious disease due to viruses or bacteria,
XX allergy, hypertension, hyperlipaemia, diabetes, cardiac disease, cerebral
XX infarction, dementia, obesity, arteriosclerosis, infertility, mental and
XX nervous diseases, cataract, progeria, or hypersensitivity to ultraviolet
XX radiation. The present sequence encodes human fibronectin 1. Bovine
XX fibronectin was used as a candidate for an actin acting substance. The
XX actin acting substance was used with transfection reagents and amplified
XX plasmid DNA in assays using transfection arrays, in the course of the
XX invention.
XX Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 6.62e-07 Length: 1929
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-581-651D-5 (1-20) x ADS17488 (1-1929)
Qy 1 IleSerLysTVrIleLeuArgTrrArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 1867 ATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCACCAGAAACCTTGGATAC 1926
RESULT 2
ADR97657
ID ADR97657 standard; DNA; 1929 BP.
XX ADR97657;
XX 02-DEC-2004 (first entry)
XX Human fibronectin 1 DNA, an actin acting substance SeqID 1.
XX human; gene; ds; transfection efficiency; actin acting substance;
XX extracellular matrix; fibronectin 1; gene introduction reagent.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 1..1929
XX /tag= a
XX /product= "Fibronectin protein"
XX WO2004079332-A2.
XX
XX 04-MAR-2003; 2003JP-00057870.
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX Miyake M, Yoshikawa T, Uchimura E, Miyake J;
XX WPI; 2004-662438/64.
XX P-PSDB; ADS17489.
XX Presenting a state of a cell, useful for diagnosing and treating a
XX disease, e.g. cancer, infectious disease, allergy, diabetes, dementia,
XX obesity, infertility, or cataract, comprises obtaining a time-lapse
XX profile of the cell.
XX Disclosure; SEQ ID NO 1; 532pp; English.
XX The specification describes a method and system for accurately presenting
XX a state of a cell. The method comprises obtaining a time-lapse profile of
XX the cell by time-lapse monitoring of a gene state associated with at
XX least one gene derived from the cell, and presenting the time-lapse
XX profile. The gene comprises a transcription control sequence, and the
XX gene state includes expression of the gene. The method and system are
XX useful presenting a state of a cell. The method can allow the elucidation
XX of key protein-protein interactions suitable for targeting by drug
XX screening protocols. The method is useful for diagnosing or treating a
XX disease, e.g. cancer, infectious disease due to viruses or bacteria,
XX allergy, hypertension, hyperlipaemia, diabetes, cardiac disease, cerebral
XX infarction, dementia, obesity, arteriosclerosis, infertility, mental and
XX nervous diseases, cataract, progeria, or hypersensitivity to ultraviolet
XX radiation. The present sequence encodes human fibronectin 1. Bovine
XX fibronectin was used as a candidate for an actin acting substance. The
XX actin acting substance was used with transfection reagents and amplified
XX plasmid DNA in assays using transfection arrays, in the course of the
XX invention.
XX Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 6.62e-07 Length: 1929
Score: 110.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-581-651D-5 (1-20) x ADR97657 (1-1929)
Qy 1 IleSerLysTVrIleLeuArgTrrArgProValSerIleProProArgAsnLeuGlyTyr 20
Db 1867 ATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCACCAGAAACCTTGGATAC 1926
RESULT 3
ADD18477
ID ADD18477 standard; DNA; 2127 BP.
XX ADD18477;
XX 15-JAN-2004 (first entry)
XX Human prostate cancer diagnosis related DNA sequence SeqID49.
XX prostate tissue; cancer diagnostic; cancer marker; prostate cancer; PCA;
XX male cancer-related death; serum biomarker; tissue biomarker; cytostatic;
XX gene therapy; prostate biopsy tissue; ANAC;
XX alpha-methylacyl-coenzyme A racemase; diagnosing cancer; cell growth;
XX human; ds.
XX Homo sapiens.
XX WO2003012067-A2.
XX
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 535.477 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-5

Perfect score: 110

Sequence: 1 ISKYLWRPVSIPPRNLGY 20

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool_h/US09581651/runat.07112005.092223.28778/app_query.fasta_1.1834
-DB=GenEmbl -QWIT=fastap -SUFFX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651.CG 1.1 9936 @runat.07112005.092223.28778 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOURY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	110	100.0	471	6	C0697541 Sequence
2	110	100.0	1929	6	C0871810 Sequence
3	110	100.0	1929	6	C0871828 Sequence
4	110	100.0	2147	6	B0137021 Polypepti

5	110	100.0	2147	6	AX003229	Sequence
6	110	100.0	2147	9	HS276395	Homo sapi
7	110	100.0	2192	9	HS276395	Homo sapi
8	110	100.0	2358	6	C0875358	Sequence
9	110	100.0	4321	9	HS276395	Homo sapi
10	110	100.0	143947	9	AC012462	Homo sapi
11	65	59.1	197121	2	AC127101	Rattus no
12	65	59.1	229287	2	AC107509	Rattus no
13	65	59.1	229287	2	AC107509	Rattus no
14	65	59.1	244185	2	AC123496	Rattus no
15	65	59.1	255929	2	AC095457	Rattus no
16	62	56.4	7085	6	A76577	Sequence 9
17	62	56.4	231764	2	AC119470	Rattus no
18	62	56.4	233445	2	AC129682	Rattus no
19	61	55.5	185651	2	AC124821	Mus muscu
20	61	55.5	204737	2	AC091456	Mus muscu
21	60	54.5	163319	2	AC110512	Mus muscu
22	60	54.5	198224	2	AC141795	Apis mell
23	59.5	54.1	287860	1	AE017265	Bacillus
24	59	53.6	543	14	AF369258	Hepatitis
25	58.5	53.2	180594	10	AC112270	Mus muscu
26	58.5	53.2	267977	2	AC151721	Mus muscu
27	58	52.7	9427	14	HPCJRNA	Hepatitis C
28	58	52.7	11235	1	AE008320	Agrobacte
29	58	52.7	11449	1	AE015266	Shigella
30	58	52.7	14357	1	AE009292	Agrobacte
31	58	52.7	34670	3	LMFL1648	Leishmani
32	58	52.7	110000	2	LMFLCHR32_09	Continuatio
33	58	52.7	140566	2	AC141532	Rattus no
34	58	52.7	215689	2	AC111687	Rattus no
35	58	52.7	292088	1	AE016986	Shigella
36	57	51.8	210	14	HPCNS3A2	Hepatitis C
37	57	51.8	3640	14	AY070174	Hepatitis
38	57	51.8	9365	14	AY587844	Hepatitis
39	57	51.8	78184	9	AC067722	Homo sapi
40	57	51.8	80938	2	AC023502	Homo sapi
41	57	51.8	106556	9	HSDJ512E2	Human DNA
42	57	51.8	109853	8	AC135417	Oryza sat
43	57	51.8	141415	2	AC008032	Homo sapi
44	57	51.8	144014	2	AC149702	Bos tauru
45	57	51.8	175916	2	AC007747	Homo sapi

ALIGNMENTS

RESULT 1	471 bp	DNA	linear	PAT 03-FEB-2004
C0697541	Sequence 42467 from Patent WO02070737.			
LOCUS	C0697541			
DEFINITION	C0697541			
ACCESSION	C0697541			
VERSION	C0697541.1	GI:42248968		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Liew C.C., Marshall, W.B. and Zhang, H.			
TITLE	Compositions and methods relating to osteoarthritis			
JOURNAL	Patent: WO 02070737-A 42467 12-SEP-2002;			
FEATURES	Chondrogene Inc. (CA)			
source	Location/Qualifiers			
	1..471			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			

ORIGIN				
Alignment Scores:				
Pred No.:	Se-08	Length:	471	
Score:	110.00	Matches:	20	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2005, 22:29:25 ; Search time 7025 Seconds
(without alignments)
11633.308 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 2147

Sequence: 1 caaacttggtggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 segs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1895.8	87.8	7777	3	CR749316 Homo sapi
2	1895.8	87.8	7885	3	CR749281 Homo sapi
3	1892.6	87.7	8121	3	CR749317 Homo sapi
4	1875.4	87.3	7501	3	BC078656 Homo sapi
5	1460.2	68.0	8315	3	AK090135 Mus muscu
6	1391.2	64.8	8329	3	AK090130 Mus muscu
7	908.6	42.3	1044	5	BC398837 BX398837
8	875.2	40.8	1052	5	BX380582 BX380582
9	794.8	37.0	943	5	BX391752 BX391752
10	784	36.5	1026	5	BX386270 BX386270
11	766.6	35.7	816	1	AU141008 AU141008
12	766.2	35.7	967	5	BX402381 BX402381
13	749	34.9	766	7	CN419594 170004245
14	748.8	34.9	1050	5	BX417945 BX417945
15	733.2	34.1	765	4	BM715855 UI-E-EJ0-
16	731	34.0	861	1	AU140971 AU140971
17	727.4	33.9	737	1	AU140993 AU140993
18	723.2	33.7	731	1	AL1706215 AU140526
19	718.2	33.5	739	1	AL706215 DKFZp686J
20	707.2	32.9	747	6	CA423317 UI-H-PE1-
21	700.2	32.6	730	1	AU140834 AU140834
22	693	32.3	695	7	CN419477 170004247
23	693	32.3	706	4	BM674187 UI-E-EJ0-
24	686.8	32.0	2107	3	AK054456 Mus muscu

25 686.6 32.0 837 1 AU140973 AU140973
26 680.8 31.7 699 1 AU140889 AU140889
27 676.6 31.5 734 1 AU140802 AU140802
28 676 31.5 734 1 AL603362 DKFZp686C
29 673.2 31.4 705 1 AU140789 AU140789
30 669 31.2 723 7 CN419622 170005319
31 664 30.9 667 1 AU140910 AU140910
32 663.2 30.9 911 5 BX398838 BX398838
33 662 30.8 738 6 CD613781 55110394J
34 660.6 30.8 670 5 BU620723 UI-H-FL1-
35 656.4 30.5 836 7 CN538822 UI-M-HS0-
36 655.4 30.5 1038 5 BX439175 BX439175
37 655.2 30.5 673 1 AU140450 AU140450
38 654 30.5 851 1 AU141024 AU141024
39 652 30.4 859 7 CF616056 AGENCOURT
40 646.4 30.1 796 1 AU140814 AU140814
41 639.8 29.8 699 1 AL706221 DKFZp686J
42 635 29.6 748 1 AU140948 AU140948
43 632.2 29.4 748 1 AU140991 AU140991
44 628.6 29.3 758 1 AU140556 AU140556
45 626.4 29.2 673 7 CN419479 170004706

ALIGNMENTS

RESULT 1
CR749316
LOCUS CR749316 7777 bp mRNA linear HTC 19-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686K08164 (from clone DKFZp686K08164).
ACCESSION CR749316
VERSION CR749316.1 GI:51476361
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 7777)
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeanger, A., Fobo G., Han, M. and Wiemann, S.
CONSRM The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMPZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseidorf/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp686K08164) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering.
http://www.rzpd.de/cgi-bin/products/cl.cgi?ClonedID=DKFZp686K08164
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES

source
Location/Qualifiers
1..7777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686K08164"
/tissue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_SfiI; host DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin 1, differentially spliced"
1..7777
/gene="DKFZp686K08164"
266..7069
/gene="DKFZp686K08164"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18171.1"

```

/db_xref="GI:51476362"
/tranlatlon="MLRGPGLGGLLLALVLCIGTAVPSTGASKSKRRQAQMVPQSPVA
VNSQKPCYDNGKHQINQWERYLGNALVCTCYGSGRGFNCKSPKASEETCFDKYT
GSTRYRGDTYRPRKDSIMWDCTCIGAGRGRICTIANRCHGGGSGYKIGDTRRPHET
GGYMLCECTCLNGKNDWTCRPIABKCFDHAAGTSYVVGETWKEPYQGMWVDCITLGE
GSGRICTCTNRKNDQTRTSYRIGTWSKDNRNLLQCICTGNRGEMWCKERTISV
QTTSSGGPFDVRAAVYQPPHPPPPYGHCVTDGVSIVVSGMQLKTOGNKQMLCT
CLNGVSCQETATQTVTGGNSGRCPLVFTYNGRTFVSCTEGRQDHLWCSTTSY
EQDQX3FCIDHTVLVQTRGNSGALCHFFPLYNHNYTDCSEGRDNWKGWGTQ
NYDADQFPCFMAAHEICTINEGVMTRIGDMQKQDMGRMKRCCVNGRGWETC
IAYSQLRDQCDIDITVNDTFHREHEEHLNCTCFQGGMRKCTCDVDQDQDSET
GTFFYIGDSWEKYHYCYCYGSGRGFNCKSPKASEETCFDKYT
SHP7QMAPOBSPKSKYLWRPKNSVGRWKEATIPGLHNSYTIKGLPGVYVGGQI
SIQYQHOEVTRPDTTSTPVSNTVCTGCTPSPVATSEVTEITASSPVSVM
VSADTVSGRPREVELESEBDEPQYLDLPSTATSVNIPDLPLGRKYLNNYQI5EDGE
QSLILSTQTTAPADPDPDQVDDTSIVVRWSRQAPITGYRIVYSPSVEGSSTEL
NLEPANSVTLSDLOPGVQYNIITIVAEENQESTPVVQOETTPKSDTVPSPRDLO
FVETDVKVITLMTTPPSAVTGYVDVIPVNLPGHGQRLPISRTTAEVGLSPGYT
YFVFAVSHGRESKPLTAQOTKLDAPTNLFQVNETDSTVLVWRTPRAQITGYRLT
VGLTRGPOPRVNVGPSVKYPLRNLOPASVYTVSLVAIKGNQESPRATGVFTTLPQ
SSIIPPVTEVETTVITVTPAPRIGPKLGVPSQGEAPREVTSDSGSIWVSGLTG
VEYVITVQLRDQERDAPVNVKVVTPSLPPTNLHLENPDTGVLTVSWERSTTPDIT
GYRITTTPTNGQNSDEEVHADQSSCTFDNLSGLEYNVSVTVXDDKRESVPSIDT
IPEVQJLDSLFDVITDSSIGLRTWPLNSSTIIIGYRTVVVAAGEGPIEFDFVDSV
GIYVTVGLEPIDGIDVITSLINGESAPTLTQOTAVPPPTDLRTNIGPDMTRVW
APPSIDLTNPLVRYSPKNEEDVAELISPSDNVAVLTNLLPGTEYVVS5VVEOH
DRVPHRSNITLNTLPTGTEYVSVIANGRESPLLIQOQSTVSDVPRDLEVVAAIP
TSLISWDAPAVTVRYRTIYGETGNSPQVEFTVPGSKSTATISGLSPKGVDTITVY
AVTGRGSPASSKPIINRYTEIDKPSQMPQVTDQNSISVWLPSSPVGTYRVTIT
PKNGPPTKTAGPDQTEMTEIGLQPTVEYVVAQNPSESPQPLVATVATVITPAP
TDLAFTOVPTSLAQWTPPNVOLTGVHVRVTPKEKTGPMEIINLAPDSSVVVSGLM
VATKYESVYALKOTLSRPAQGVVTVLENVSPRRARVDTATETITISWRTXETIT
TGFQDVPANGQPIQKIIIPDVRVYITIGLQPDYDKIYLTILNDNARSPPVIDA
STADAPNSRFLATPNTSLNSWQPPRATYIIKYKPGSPREVPVPRPGVTE
ATITGLEPQYFVYVIALKNKQSEPLIGRTKQGEALSTQITISWAPQDTSYIIS
CHPTVDEEPLQYFVYVIALKNKQSEPLIGRTKQGEALSTQITISWAPQDTSYIIS
VNEGLNOPDSDPDPYVTVSHVAVGDEWRESGFKLLCOCLGFGSHFRCDSDRSWC
HDNGVNYKIGKMBQDQNGNGOMSCITCLNGKBEKFDKPDHEATCYDQKTYHVEQW
KEYLGAICSCCTCFGQGRWRCNCRPGCEBTPBEGTTPGTSQSNQSORHQRTNTVNC
PIECFMPJDVQADREDSRE"

```

ORIGIN

```

Query Match      87.8%; Score 1885.8; DB 3; Length 7777;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 7; Indels 45; Gaps 1;

Qy 1 CAAACTTGGTGCAACTTGCCTCCCGGTGCGGGCGTCTCTCCCCACCGTCTCAACATGC 60
Db 210 CAAACTTGGTGCAACTTGCCTCCCGGTGCGGGCGTCTCTCCCCACCGTCTCAACATGC 269

Qy 61 TTAGGGGTGCGGGGCGGGGTGCTGCTGCTGGCCGTCCAGTGCCTGGGACAGGGTGC 120
Db 270 TTAGGGGTGCGGGGCGGGGTGCTGCTGCTGGCCGTCCAGTGCCTGGGACAGGGTGC 329

Qy 121 CCTCACCGGAGCCTCGAAGACAGAGAGCGGCTCAGCAATGCTTCAGCCCGCAGTCCC 180
Db 330 CCTCACCGGAGCCTCGAAGACAGAGAGCGGCTCAGCAATGCTTCAGCCCGCAGTCCC 389

Qy 181 CGGTGGCTGTTCAGTCAAAAGCAAGCCCGGTGTTATGACAAATGGAAACACATATCAGATAA 240
Db 390 CGGTGGCTGTTCAGTCAAAAGCAAGCCCGGTGTTATGACAAATGGAAACACATATCAGATAA 449

Qy 241 ATCAACAGTGGAGGGAGCTACCTAGGCAATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
Db 450 ATCAACAGTGGAGGGAGCTACCTAGGCAATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 509

Qy 301 GCCGAGGTTTAACTCGGAGAGTAAACCTGAAGCTTGAAGAGACTTGTCTTGAAGAGTACA 360
Db 510 GCCGAGGTTTAACTCGGAGAGTAAACCTGAAGCTTGAAGAGACTTGTCTTGAAGAGTACA 569

Qy 361 CTGGGAACA CTTTACCGATGGGTGACACTTATGAGCGTCTTAAAGACTCCCATGATCTGGG 420

```

```

Db 570 CTGGGAACA CACTTACCAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCCATGATCTGGG 629
Qy 421 ACTGTACTCTCATCGGGGCTGGGCGAGGAGATAAGCTGTACCATCGCAAAACGCTGCC 480
Db 630 ACTGTACTCTCATCGGGGCTGGGCGAGGAGATAAGCTGTACCATCGCAAAACGCTGCC 689
Qy 481 ATGAAGGGGGTCAAGTCTCTCAAGAATTGGTGACACCTCGAGGAGACCAATGAGACTGGTG 540
Db 690 ATGAAGGGGGTCAAGTCTCTCAAGAATTGGTGACACCTCGAGGAGACCAATGAGACTGGTG 749
Qy 541 GTTACATGTTAGAGTGTGTCTTGGTAAATGGAAAGAGAGATGACCTGCAAGCCCA 600
Db 750 GTTACATGTTAGAGTGTGTCTTGGTAAATGGAAAGAGAGATGACCTGCAAGCCCA 809
Qy 601 TAGCTGAGAAAGTGTGTTTGTATCATCTCTCGGACTTCTTATGTGTCGGAAGAAACGTGGG 660
Db 810 TAGCTGAGAAAGTGTGTTTGTATCATCTCTCGGACTTCTTATGTGTCGGAAGAAACGTGGG 869
Qy 661 AGAAGCCCTTACCAAGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 870 AGAAGCCCTTACCAAGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 929
Qy 721 GCATCACTTGCATCTTCTAGAAAATAGATGCAACGATCAGGACACAGACATCCTATAGAA 780
Db 930 GCATCACTTGCATCTTCTAGAAAATAGATGCAACGATCAGGACACAGACATCCTATAGAA 989
Qy 781 TTGAGACACCTGAGCAAGAGGATTAATCGAGAAACCTGCTCCAGTGCATCTGCACAG 840
Db 990 TTGAGACACCTGAGCAAGAGGATTAATCGAGAAACCTGCTCCAGTGCATCTGCACAG 1049
Qy 841 GCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGTCAGACCAATCGAGCG 900
Db 1050 GCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGTCAGACCAATCGAGCG 1109
Qy 901 GATCTGGCCCTTTCACCGATGTTTCTGTCAGCTGTTTACCAACCGCAGCTCACCCGAGC 960
Db 1110 GATCTGGCCCTTTCACCGATGTTTCTGTCAGCTGTTTACCAACCGCAGCTCACCCGAGC 1169
Qy 961 CTCTCTCCTATGGCCACTGTGTACACAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 1170 CTCTCTCCTATGGCCACTGTGTACACAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1229
Qy 1021 GGCTGAAGACACAAAGGAAATAAGCAATGCTTTCGACGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1080
Db 1230 GGCTGAAGACACAAAGGAAATAAGCAATGCTTTCGACGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1289
Qy 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACCGTGGCACTCAAAATGGAGAGCCATGTGTCT 1140
Db 1290 GCCAAGAGACAGCTGTAAACCCAGACTTACCGTGGCACTCAAAATGGAGAGCCATGTGTCT 1349
Qy 1141 TACCATTACCTCAACAGCAGGAC----- 1165
Db 1350 TACCATTACCTCAACAGCAGGAC----- 1409
Qy 1166 -----GGACAGACACAACTTCGAATTATGAGAGGACAGCAAAATATCTTTCTGCA 1215
Db 1410 GACATCTTTGGTGCAGCAGCAACTTCGAATTATGAGAGGACAGCAAAATATCTTTCTGCA 1469
Qy 1216 CAGACACACTGTTTGGTTCAGACTCGAGAGGAAATTCCAATGGTGGCTGTGCCACT 1275
Db 1470 CAGACACACTGTTTGGTTCAGACTCGAGAGGAAATTCCAATGGTGGCTGTGCCACT 1529
Qy 1276 TCCCTCTCTATACAAACACCAATTAACACTGATGATGATGATGATGATGATGATGATGATGATGATG 1335
Db 1530 TCCCTCTCTATACAAACACCAATTAACACTGATGATGATGATGATGATGATGATGATGATGATG 1589
Qy 1336 ACATGAAGTGGTGGGACCAACAGAACTATGATGATGATGATGATGATGATGATGATGATGATG 1395
Db 1590 ACATGAAGTGGTGGGACCAACAGAACTATGATGATGATGATGATGATGATGATGATGATGATG 1649
Qy 1396 CCATGGTGGCCACGAGGAAATCTGCACCAACCAATGAGGGGTGATGATGATGATGATGATGATG 1455
Db 1650 CCATGGTGGCCACGAGGAAATCTGCACCAACCAATGAGGGGTGATGATGATGATGATGATGATG 1709

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 11, 2005, 01:52:51 ; Search time 1727 Seconds
(without alignments)
10281.009 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaactgtggaacttgc.....aaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2120.6	98.8	2127	22	US-10-956-157-4288
2	2119.8	98.7	2443	20	Sequence 4288, Ap
3	2119.8	98.7	2443	22	Sequence 70, Appl
4	2119	98.7	2127	17	US-10-741-601-70
5	2119	98.7	2127	24	Sequence 238, App
					Sequence 49, Appl
					Sequence 49, Appl

6	2064.8	96.2	2488	20	US-10-741-601-75	Sequence 75, Appl
7	2064.8	96.2	2488	22	US-10-741-600-244	Sequence 244, App
8	1975.8	92.0	4295	18	US-10-144-194A-51	Sequence 51, Appl
9	1975.8	92.0	4295	18	US-10-491-566-51	Sequence 51, Appl
10	1890.6	88.1	8027	21	US-10-447-161-8	Sequence 8, Appl1
11	1890.6	88.1	8027	20	US-10-734-564-27	Sequence 27, Appl
12	1890.6	88.1	8027	22	US-10-852-335A-53	Sequence 53, Appl
13	1890.6	88.1	8027	24	US-10-287-436A-81	Sequence 81, Appl
14	1890.6	88.1	8815	22	US-10-868-577A-62	Sequence 62, Appl
15	1890.6	88.1	8815	22	US-10-868-549-21	Sequence 21, Appl
16	1889.8	88.0	6510	20	US-10-741-601-72	Sequence 72, Appl
17	1889.8	88.0	6510	20	US-10-741-600-241	Sequence 241, App
18	1889.8	88.0	7823	20	US-10-741-601-77	Sequence 77, Appl
19	1889.8	88.0	7823	22	US-10-741-600-245	Sequence 245, App
20	1889.8	88.0	7848	20	US-10-741-601-78	Sequence 78, Appl
21	1889.8	88.0	7848	22	US-10-741-600-246	Sequence 246, App
22	1889.8	88.0	7935	20	US-10-741-601-74	Sequence 74, Appl
23	1889.8	88.0	7935	22	US-10-741-600-240	Sequence 240, App
24	1889.8	88.0	7935	20	US-10-741-601-81	Sequence 81, Appl
25	1889.8	88.0	7959	22	US-10-741-600-249	Sequence 249, App
26	1889.8	88.0	8013	20	US-10-741-601-71	Sequence 71, Appl
27	1889.8	88.0	8013	22	US-10-741-600-242	Sequence 242, App
28	1889.8	88.0	8155	20	US-10-741-601-79	Sequence 79, Appl
29	1889.8	88.0	8155	22	US-10-741-600-247	Sequence 247, App
30	1889.8	88.0	8226	20	US-10-741-601-69	Sequence 69, Appl
31	1889.8	88.0	8226	22	US-10-741-600-237	Sequence 237, App
32	1889.8	88.0	8278	20	US-10-741-601-82	Sequence 82, Appl
33	1889.8	88.0	8278	22	US-10-741-600-250	Sequence 250, App
34	1889.8	88.0	8332	20	US-10-741-601-73	Sequence 73, Appl
35	1889.8	88.0	8332	22	US-10-741-600-239	Sequence 239, App
36	1889.8	88.0	8371	20	US-10-741-601-76	Sequence 76, Appl
37	1889.8	88.0	8371	22	US-10-741-600-243	Sequence 243, App
38	1887.4	87.9	7795	16	US-10-084-817-2	Sequence 2, Appl1
39	1887.4	87.9	7867	14	US-10-098-841-6	Sequence 6, Appl1
40	1887.4	87.9	8044	17	US-10-240-965-121	Sequence 121, App
41	1887.4	87.9	8044	24	US-10-765-700-135	Sequence 135, App
42	1887.4	87.9	8062	14	US-10-098-841-5	Sequence 5, Appl1
43	1887.4	87.9	8137	14	US-10-098-841-8	Sequence 8, Appl1
44	1887.4	87.9	8230	14	US-10-098-841-7	Sequence 7, Appl1
45	1839.6	85.7	7361	19	US-10-236-392-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1

US-10-956-157-4288
; Sequence 4288, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4288
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-4288

Query Match 98.8%; Score 2120.6; DB 22; Length 2127;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2123; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CAAACTGTGGCAACTTGCTCCCGGTGGGGGGTCTCTCCCGCACCGTCTCAACATGC 60

Db 1 CAAACTGTGGCAACTTGCTCCCGGTGGGGGGTCTCTCTCCCGCACCGTCTCAACATGC 60

Result No.	Query	Score	Match	Length	DB	ID	Description
C	1	1890.6	88.1	8232	6	US-10-821-234-693	Sequence 693, Appl
	2	772.2	36.0	777	6	US-10-631-558-3	Sequence 3, Appli
	3	330	15.4	1541	6	US-10-631-558-9	Sequence 9, Appli
	4	330	15.4	2096	6	US-10-631-558-12	Sequence 12, Appl
	5	312	14.5	1661	6	US-10-631-558-10	Sequence 10, Appl
	6	232.8	10.8	1782	6	US-10-631-558-11	Sequence 11, Appl
	7	162.2	7.6	2154	6	US-10-750-185-25146	Sequence 25146, A
	8	73.8	3.4	2334	8	US-11-112-944-22	Sequence 22, Appl
	9	71.8	3.3	3069	6	US-10-909-123-824	Sequence 82, Appl
c	10	71.8	3.3	3515	6	US-10-821-234-65	Sequence 65, Appl
	11	48.2	2.2	1672	7	US-11-102-240-17	Sequence 17, Appl
	12	37.6	1.8	257	6	US-10-623-155-48	Sequence 48, Appl
	13	35.4	1.6	2822	6	US-10-131-826A-305	Sequence 305, App
	14	34.8	1.6	1113	6	US-10-750-185-41675	Sequence 41675, A
	15	34.8	1.6	2140	7	US-11-087-221-17	Sequence 17, Appl
	16	34.8	1.6	2281	7	US-11-087-227-19	Sequence 19, Appl
	17	34.8	1.6	149419	7	US-11-112-908-49	Sequence 49, Appl
	18	34.8	1.6	166111	7	US-11-112-908-47	Sequence 47, Appl
19	34.6	1.6	1511	7	US-11-102-978-6	Sequence 6, Appli	
20	34.6	1.6	4053	6	US-10-131-828A-351	Sequence 351, App	
21	34.6	1.6	319608	7	US-11-145-703-1	Sequence 1, Appli	
22	34.4	1.6	2745	6	US-10-517-544-76	Sequence 76, Appl	
23	34.2	1.6	1828	6	US-10-510-386-99	Sequence 99, Appl	

Qy 301 GCCGAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAATACA 360
 Db 7721 GCCGAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAATACA 7662
 Qy 361 CTGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTTAAGAGCTTCCATGATCTGGG 420
 Db 7661 CTGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTTAAGAGCTTCCATGATCTGGG 7602
 Qy 421 ACTGTACTGATCGGGCTGGGCGAGGGAGAAATAAGCTGTACCATGCCAACCGCTGCC 480
 Db 7601 ACTGTACTGATCGGGCTGGGCGAGGGAGAAATAAGCTGTACCATGCCAACCGCTGCC 7542
 Qy 481 ATGAAGGGGGTCACTCTTACAAAGATTGGTGACACTGGAGGAGACCATGAGACTGGTG 540
 Db 7541 ATGAAGGGGGTCACTCTTACAAAGATTGGTGACACTGGAGGAGACCATGAGACTGGTG 7482
 Qy 541 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGAGAGAAATGGACCTGCAAGCCCA 600
 Db 7481 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGAGAGAAATGGACCTGCAAGCCCA 7422
 Qy 601 TAGCTGAGAAGTGTGATCATGCTGTGGGACTTCTTATGTGCTGGGAGAGCGAGCTGGG 660
 Db 7421 TAGCTGAGAAGTGTGATCATGCTGTGGGACTTCTTATGTGCTGGGAGAGCGAGCTGGG 7362
 Qy 661 AGAAGCCCTACCAAGGCTGGATGATGGTGAATTTGATCTGGGAGAGCGAGCTGGG 720
 Db 7361 AGAAGCCCTACCAAGGCTGGATGATGGTGAATTTGATCTGGGAGAGCGAGCTGGG 7302
 Qy 721 GCATCATTGCACTTTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTATAGAA 780
 Db 7301 GCATCATTGCACTTTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTATAGAA 7242
 Qy 781 TTGAGACACCTGGAGCAAGAGATATCAGGAAACCTGCTCAGTGCATCTGCACAG 840
 Db 7241 TTGAGACACCTGGAGCAAGAGATATCAGGAAACCTGCTCAGTGCATCTGCACAG 7182
 Qy 841 GCAACGGCCGAGGAGTGGAAAGTGTGAGAGGACACACCTCTGTGAGACACATCTGAGCG 900
 Db 7181 GCAACGGCCGAGGAGTGGAAAGTGTGAGAGGACACACCTCTGTGAGACACATCTGAGCG 7122
 Qy 901 GATCTGGCCCTTCAACCGATGTTGTCAGCTGTTTACCAACCGAGCTTCAACCCCGAGC 960
 Db 7121 GATCTGGCCCTTCAACCGATGTTGTCAGCTGTTTACCAACCGAGCTTCAACCCCGAGC 7062
 Qy 961 CTCCTCCCTATGGCCACTGTGTACAGACAGTGGTGTGCTCTCTGTGGGATGCAGT 1020
 Db 7061 CTCCTCCCTATGGCCACTGTGTACAGACAGTGGTGTGCTCTCTGTGGGATGCAGT 7002
 Qy 1021 GGCTGAACACACAAAGGAATAGCAAAATGCTTTGACAGTGTGCTGGCAACCGAGTCACT 1080
 Db 7001 GGCTGAACACACAAAGGAATAGCAAAATGCTTTGACAGTGTGCTGGCAACCGAGTCACT 6942
 Qy 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCACTCAAAATGGAGAGCAATGTGTCT 1140
 Db 6941 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCACTCAAAATGGAGAGCAATGTGTCT 6882
 Qy 1141 TACATTCACCTAACAGCAGGAGC----- 1165
 Db 6881 TACATTCACCTAACAGCAGGAGCTTCTACTCTGCACCAACGGAAGGGCGACAGGAGC 6822
 Qy 1166 -----GGACAGACAACTTCGAATTTAGCAGGAGCAGCAAAATACTCTTTCTGCA 1215
 Db 6821 GACATCTTTGGTGCAGCAAACTTCGAATTTAGCAGGAGCAGCAAAATACTCTTTCTGCA 6762
 Qy 1216 CAGACACACTGTTTTGGTTTCAGACTCGAGGAGGAAATTCGAATGGTGCCTTTGTGCCACT 1275
 Db 6761 CAGACACACTGTTTTGGTTTCAGACTCGAGGAGGAAATTCGAATGGTGCCTTTGTGCCACT 6702
 Qy 1276 TCCCTCTTCTATACAAACCAACCAATTTACATGATTGCACTTCTGAGGGCAGAGAGACA 1335
 Db 6701 TCCCTCTTCTATACAAACCAACCAATTTACATGATTGCACTTCTGAGGGCAGAGAGACA 6642

Qy 1336 ACATGAAGTGGTGTGGGACACACAGAACTATGATCCGACACAGAACTTTGGGTCTGCC 1395
 Db 6641 ACATGAAGTGGTGTGGGACACACAGAACTATGATCCGACACAGAACTTTGGGTCTGCC 6582
 Qy 1396 CCATGGTCCCAAGAGGAAATCTGCACAAACCAATGAAGGGGTCACTACCGCATTTGGAG 1455
 Db 6581 CCATGGTCCCAAGAGGAAATCTGCACAAACCAATGAAGGGGTCACTACCGCATTTGGAG 6522
 Qy 1456 ATCAGTGGGATAAGCAGCATGACATGGTGCACATGATGAGTGCACGTGTGTGGGAATG 1515
 Db 6521 ATCAGTGGGATAAGCAGCATGACATGGTGCACATGATGAGTGCACGTGTGTGGGAATG 6462
 Qy 1516 GTCGTGGGAAATGGACATGCAATTCCTTACTCGCAGCTTCAGATCAGTGCATTTGATG 1575
 Db 6461 GTCGTGGGAAATGGACATGCAATTCCTTACTCGCAGCTTCAGATCAGTGCATTTGATG 6402
 Qy 1576 ACATCACTTCAATGTGAACGACACATTCACAAAGCGTCAATGAAGGGGCGACATGCTGA 1635
 Db 6401 ACATCACTTCAATGTGAACGACACATTCACAAAGCGTCAATGAAGGGGCGACATGCTGA 6342
 Qy 1636 ACTGTACATCTTCGGTTCAGGGTCGGGCGAGGTGGAAGTGTGATCCCGTCCGACCAATGCC 1695
 Db 6341 ACTGTACATCTTCGGTTCAGGGTCGGGCGAGGTGGAAGTGTGATCCCGTCCGACCAATGCC 6282
 Qy 1696 AGGATTCAGAGACTGGGACGCTTTTATCAAAATTCGAGATTTCATGGGAGAAATGATGCAATG 1755
 Db 6281 AGGATTCAGAGACTGGGACGCTTTTATCAAAATTCGAGATTTCATGGGAGAAATGATGCAATG 6222
 Qy 1756 GTGTCAAGTACCAGTGTCTACTGTATGGCCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 1815
 Db 6221 GTGTCAAGTACCAGTGTCTACTGTATGGCCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 6162
 Qy 1816 TACAGACTTCAAGCTCAAGTGGTCTCTCGAAGTATTTATCAGTGAAGTCCGAGTCCGAGTC 1875
 Db 6161 TACAGACTTCAAGCTCAAGTGGTCTCTCGAAGTATTTATCAGTGAAGTCCGAGTCCGAGTC 6102
 Qy 1876 AGCCCAACTCCCAACCCATCCAGTGGAAATGCACACAGCCATCTCACATTTCCAAGTACA 1935
 Db 6101 AGCCCAACTCCCAACCCATCCAGTGGAAATGCACACAGCCATCTCACATTTCCAAGTACA 6042
 Qy 1936 TTCTCAGTGGAGACCT 1952
 Db 6041 TTCTCAGTGGAGACCT 6025

RESULT 2
 US-10-631-558-3
 ; Sequence 3, Application US/10631558
 ; Publication No. US20050260598A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kumar, Rajesh
 ; APPLICANT: Sahni, Girish
 ; APPLICANT: Roy, Chait
 ; APPLICANT: Rajagopal, Kammar
 ; APPLICANT: Nihalani, Deepak
 ; APPLICANT: Sundaram, Vasudha
 ; APPLICANT: Yadav, Mahavir
 ; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
 ; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
 ; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 07064-009002
 ; CURRENT APPLICATION NUMBER: US/10/631,558
 ; CURRENT FILING DATE: 2003-07-31
 ; PRIOR APPLICATION NUMBER: US/09/940,235
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 09/471,349
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
 ; PRIOR FILING DATE: 1998-12-24
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 21:28:04 ; Search time 1571 Seconds
(without alignments)
11301.308 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 2147

Sequence: 1 caaactgtggaacttgc.....aaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA Main.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2120.6	98.8	2127	9	US-10-956-157-4288
2	2119.8	98.7	2443	7	US-10-741-601-70
3	2119.8	98.7	2443	8	US-10-741-600-238
4	2119	98.7	2127	6	US-10-210-320-49
5	2119	98.7	2127	9	US-10-909-035-49
6	2064.8	96.2	2488	7	US-10-741-601-75
7	2064.8	96.2	2488	8	US-10-741-600-244
8	1975.8	92.0	4295	6	US-10-144-194A-51
9	1975.8	92.0	4295	8	US-10-491-566-51
10	1890.6	88.1	8027	7	US-10-447-161-8
11	1890.6	88.1	8027	7	US-10-734-564-27
12	1890.6	88.1	8027	9	US-10-852-335A-53
13	1890.6	88.1	8027	9	US-10-287-436A-81
14	1890.6	88.1	8815	8	US-10-868-577A-62
15	1890.6	88.1	8815	8	US-10-868-549-21
16	1889.8	88.0	6510	7	US-10-741-601-72
17	1889.8	88.0	6510	8	US-10-741-600-241
18	1889.8	88.0	7823	7	US-10-741-601-77
19	1889.8	88.0	7823	8	US-10-741-600-245
20	1889.8	88.0	7848	7	US-10-741-601-78
21	1889.8	88.0	7848	8	US-10-741-600-246
22	1889.8	88.0	7935	7	US-10-741-601-74
23	1889.8	88.0	7935	8	US-10-741-600-240

24	1889.8	88.0	7959	7	US-10-741-601-81	Sequence 81, Appl
25	1889.8	88.0	7959	8	US-10-741-600-249	Sequence 249, App
26	1889.8	88.0	8013	7	US-10-741-601-71	Sequence 71, Appl
27	1889.8	88.0	8013	8	US-10-741-600-242	Sequence 242, App
28	1889.8	88.0	8155	7	US-10-741-601-79	Sequence 79, Appl
29	1889.8	88.0	8155	8	US-10-741-600-247	Sequence 247, App
30	1889.8	88.0	8226	7	US-10-741-601-69	Sequence 69, Appl
31	1889.8	88.0	8226	8	US-10-741-600-237	Sequence 237, App
32	1889.8	88.0	8278	7	US-10-741-601-82	Sequence 82, Appl
33	1889.8	88.0	8278	8	US-10-741-600-250	Sequence 250, App
34	1889.8	88.0	8332	7	US-10-741-601-73	Sequence 73, Appl
35	1889.8	88.0	8332	8	US-10-741-600-239	Sequence 239, App
36	1889.8	88.0	8371	7	US-10-741-601-76	Sequence 76, Appl
37	1889.8	88.0	8371	8	US-10-741-600-243	Sequence 243, App
38	1887.4	87.9	7795	5	US-10-084-817-2	Sequence 2, Appl
39	1887.4	87.9	7867	5	US-10-098-841-6	Sequence 6, Appl
40	1887.4	87.9	8044	6	US-10-240-965-121	Sequence 121, App
41	1887.4	87.9	8044	9	US-10-765-700-135	Sequence 135, App
42	1887.4	87.9	8062	5	US-10-098-841-5	Sequence 5, Appl
43	1887.4	87.9	8137	5	US-10-098-841-8	Sequence 8, Appl
44	1887.4	87.9	8230	5	US-10-098-841-7	Sequence 7, Appl
45	1839.6	85.7	7361	7	US-10-236-392-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-956-157-4288

; Sequence 4288, Application US/10956157

; Publication No. US20050118625A1

; GENERAL INFORMATION:

; APPLICANT: Wyeth

; APPLICANT: Mounts, William

; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH

; FILE REFERENCE: 031896-043000 (AM 101081)

; CURRENT APPLICATION NUMBER: US/10/956,157

; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 4288

; LENGTH: 2127

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-956-157-4288

Query Match 98.8%; Score 2120.6; DB 9; Length 2127;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2123; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1	CAAACTTGGTGGCAACTTGGCTCCCGGTGGGGGGTCTCTCCCGCCACCGTCTCAACATGC	60
Db	1	CAAACTTGGTGGCAACTTGGCTCCCGGTGGGGGGTCTCTCCCGCCACCGTCTCAACATGC	60
Qy	61	TTAGGGGTCCGGGGCCGGGCTGCTGCTGTGGCCGTCAGTGTCTGGGACAGCGGTGC	120
Db	61	TTAGGGGTCCGGGGCCGGGCTGCTGCTGTGGCCGTCAGTGTCTGGGACAGCGGTGC	120
Qy	121	CTTCCACGGGAGCCTCGAAGAGCAAGAGGAGGCTCAGCAATGGTTTCAGCCCCCAGTCCC	180
Db	121	CTTCCACGGGAGCCTCGAAGAGCAAGAGGAGGCTCAGCAATGGTTTCAGCCCCCAGTCCC	180
Qy	181	CGGTGGCTGTGAGTCAAAAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAA	240
Db	181	CGGTGGCTGTGAGTCAAAAGCAAGCCCGTTGTTATGCAATGGAACACATATCAGATAA	240
Qy	241	ATCAACAGTGGGACCGGACCTACTAGGCAATCGTTGGTTGTACTTGTATGGAGAA	300
Db	241	ATCAACAGTGGGACCGGACCTACTAGGCAATCGTTGGTTGTACTTGTATGGAGAA	300
Qy	301	GCCGAGGTTTAACTGCGAGAGTAAACCTAGAGAGACTTGTCTTGGACAGTACA	360

```
Db 301 GCCAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACA 360
Qy 361 CTGGGAACACTTACCGAGTGGTGACACTTATGAGCGTCCTTAAAGACTCCCATGATCTGGG 420
Db 361 CTGGGAACACTTACCGAGTGGTGACACTTATGAGCGTCCTTAAAGACTCCCATGATCTGGG 420
Qy 421 ACTGTACTGTGCATCGGGCTGGGCGAGGGAGAAATAAGCTGTACCATCGCAAAACGCTGCC 480
Db 421 ACTGTACTGTGCATCGGGCTGGGCGAGGGAGAAATAAGCTGTACCATCGCAAAACGCTGCC 480
Qy 481 ATGAAGGGGTGAGTCTCTAAGAATTGGTGAACCTGTGAGAGAGACCATGAGAGCTGGTG 540
Db 481 ATGAAGGGGTGAGTCTCTAAGAATTGGTGAACCTGTGAGAGAGACCATGAGAGCTGGTG 540
Qy 541 GTTACATGTTAGAGTGTGTCTTGTGTAATGAAAGAGAGAAATGGAACCTGCAAGGCCA 600
Db 541 GTTACATGTTAGAGTGTGTCTTGTGTAATGAAAGAGAGAAATGGAACCTGCAAGGCCA 600
Qy 601 TAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCATATGTTGGTGGGAAACGTTGG 660
Db 601 TAGCTGAGAAAGTGTGTTGATCATGCTGCTGGGACTTCCATATGTTGGTGGGAAACGTTGG 660
Qy 661 AGAAGCCCTACCAAGGCTGGATGATGTGATGATTGTACTTGCTGGGAGAGGCGAGCGAC 720
Db 661 AGAAGCCCTACCAAGGCTGGATGATGTGATGATTGTACTTGCTGGGAGAGGCGAGCGAC 720
Qy 721 GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCTTATAGAA 780
Db 721 GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCTTATAGAA 780
Qy 781 TTGGAGACACCTGGAGCAAGAGGATATCAGGAAACCTGCTCCAGTGCATCTGCACAG 840
Db 781 TTGGAGACACCTGGAGCAAGAGGATATCAGGAAACCTGCTCCAGTGCATCTGCACAG 840
Qy 841 GCAAGCGCGAGGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGAGACACCATCGAGCG 900
Db 841 GCAAGCGCGAGGAGAGTGGAAAGTGTGAGAGGACACACCTCTGTGAGACACCATCGAGCG 900
Qy 901 GATCTGGCCCTTCCACGATGTTCTGTCAGCTGTTTACCAACCGCAGCTCCACCCCGAGC 960
Db 901 GATCTGGCCCTTCCACGATGTTCTGTCAGCTGTTTACCAACCGCAGCTCCACCCCGAGC 960
Qy 961 CTCCTCCCTATGGCCACTGTGTACAGACAGTGGTGTCTACTCTGTGGGGATGCAGT 1020
Db 961 CTCCTCCCTATGGCCACTGTGTACAGACAGTGGTGTCTACTCTGTGGGGATGCAGT 1020
Qy 1021 GGCTGAAGACACAAGGAAATAGCAATGCTTTGCAAGTGGTGGTGGTGGTGGTGGTGGTGG 1080
Db 1021 GGCTGAAGACACAAGGAAATAGCAATGCTTTGCAAGTGGTGGTGGTGGTGGTGGTGGTGG 1080
Qy 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCACTCAAAATGGAGAGCCATGTGTCT 1140
Db 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCACTCAAAATGGAGAGCCATGTGTCT 1140
Qy 1141 TACCATTACCTACACGACAGGACGGACAGCAACACTTTCGAAATTTATGAGCAGGACCCAGA 1200
Db 1141 TACCATTACCTACACGAGGAGCTGCAGCAACACTTCGAAATTTATGAGCAGGACCCAGA 1200
Qy 1201 AATACTCTTTTGCACAGACACACTGTTTGGTTCAGACTCGAGGAGGAAATTCCTCAATG 1260
Db 1201 AATACTCTTTTGCACAGACACACTGTTTGGTTCAGACTCGAGGAGGAAATTCCTCAATG 1260
Qy 1261 GTGCTTGTGCGACTTCCCTTCTATACACCAACCAATTTACACTGATTTGCACTTCTG 1320
Db 1261 GTGCTTGTGCGACTTCCCTTCTATACACCAACCAATTTACACTGATTTGCACTTCTG 1320
Qy 1321 AGGGCAGAGAGACAACTAAGTGGTGGTGGACCAACAGAACTATGATGCCGACCCAGA 1380
Db 1321 AGGGCAGAGAGACAACTAAGTGGTGGTGGACCAACAGAACTATGATGCCGACCCAGA 1380
Qy 1381 AGTTTGGGTTCTGCCCTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1440
Db 1381 AGTTTGGGTTCTGCCCTATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1440
```

```
Qy 1441 TGTAACCCGATTTGGAGATCAGTGGGATGAAGCAGCATGATGGGTACATGATGAGGTGCA 1500
Db 1441 TGTAACCCGATTTGGAGATCAGTGGGATGAAGCAGCATGATGGGTACATGATGAGGTGCA 1500
Qy 1501 CGTGTGTTGGGAATGTCGTGGGGAATGGACATGCATTCCTACTCGCAGCTTCGAGATC 1560
Db 1501 CGTGTGTTGGGAATGTCGTGGGGAATGGACATGCATTCCTACTCGCAGCTTCGAGATC 1560
Qy 1561 AGTGCAATTTGATGATGATCATTACAAATGTGAACGACATCTTCAAAATTTGGAGATTCATGAG 1620
Db 1561 AGTGCAATTTGATGATGATCATTACAAATGTGAACGACATCTTCAAAATTTGGAGATTCATGAG 1620
Qy 1621 AGGGGCACATGCTGAACTGTACATGCTTCGGTTCAGGCTCGGGGAGGTGGAAAGTGTGATC 1680
Db 1621 AGGGGCACATGCTGAACTGTACATGCTTCGGTTCAGGCTCGGGGAGGTGGAAAGTGTGATC 1680
Qy 1681 CCGTGCAGCAATGCGCAGGATTCAGAGACTGGGAGCTGCTATGCGCTGGCGATTCGGGAGT 1740
Db 1681 CCGTGCAGCAATGCGCAGGATTCAGAGACTGGGAGCTGCTATGCGCTGGCGATTCGGGAGT 1740
Qy 1741 AGAAGTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 AGAAGTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Qy 1801 GGCAATTCGCAACCTTTTACAGACCTATCCAAAGCTCAAGTGGTCTGTCTGTAAGTATTTATCA 1860
Db 1801 GGCAATTCGCAACCTTTTACAGACCTATCCAAAGCTCAAGTGGTCTGTCTGTAAGTATTTATCA 1860
Qy 1861 CTGAGACTCGAGTCCAGCCAACTCCCAACCCCATCCAGTGGAAATGCAACCAAGCATCTC 1920
Db 1861 CTGAGACTCGAGTCCAGCCAACTCCCAACCCCATCCAGTGGAAATGCAACCAAGCATCTC 1920
Qy 1921 ACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCCAGAAACCTTGGAT 1980
Db 1921 ACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCCAGAAACCTTGGAT 1980
Qy 1981 ACTGAGTCTCCTAATCTTATCAATCTGTGATGTTCTTTTTTCCAGCTTTTGAGCCAA 2040
Db 1981 ACTGAGTCTCCTAATCTTATCAATCTGTGATGTTCTTTTTTCCAGCTTTTGAGCCAA 2040
Qy 2041 CAATCTGATTAATCTTCTATAGCAATTTACTATATTTGTTAGTGAACAAATATG 2100
Db 2041 CAATCTGATTAATCTTCTATAGCAATTTACTATATTTGTTAGTGAACAAATATG 2100
Qy 2101 TGGTCAATTAATTTGACTTTGAGACTG 2127
Db 2101 TGGTCAATTAATTTGACTTTGAGACTG 2127
```

RESULT 2

```
US-10-741-601-70
; Sequence 70, Application US/10741601
; Publication No. US20040168519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-70
```

```
Query Match 98.7%; Score 2119.8; DB 7; Length 2443;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2121; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 17:53:25 ; Search time 7771 Seconds
(without alignments)
12926.504 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaactgtgtggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_hic.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_est7.*
9: gb_est8.*
10: gb_gsa2.*
11: gb_gsa3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1889	88.0	7868	4 HSM806653	BX640608 Homo sapi
2	1889	88.0	8411	4 HSM806992	BX640875 Homo sapi
3	1885.8	87.8	7777	4 CR749316	CR749316 Homo sapi
4	1885.8	87.8	7885	4 CR749281	CR749281 Homo sapi
5	1882.6	87.7	8121	4 CR749317	CR749317 Homo sapi
6	1875.4	87.3	7501	4 BC078656	BC078656 Homo sapi
7	1875.4	87.3	7501	4 BC100030	BC100030 Homo sapi
8	1460.2	68.0	8315	4 AK090135	AK090135 Mus muscu
9	1391.2	64.8	8329	4 AK090130	AK090130 Mus muscu
10	1335.4	62.2	7434	11 DQ039102	DQ039102 Homo sapi
11	1034.6	48.2	7434	11 DQ039103	DQ039103 Pan trogl
C 12	908.6	42.3	1044	5 BX398837	BX398837 BX398837
C 13	875.2	40.8	1052	5 BX380582	BX380582 BX380582
C 14	794.8	37.0	943	5 BX391752	BX391752 BX391752
15	784	36.5	1026	5 BX386270	BX386270 BX386270
16	766.6	35.7	816	1 AU141008	AU141008 AU141008
C 17	766.2	35.7	967	5 BX402381	BX402381 BX402381
18	749	34.9	766	5 CN419594	CN419594 170004245
19	748.8	34.9	1050	5 BX417945	BX417945 BX417945
20	733.2	34.1	765	3 BM715855	BM715855 UI-E-E30-
21	731	34.0	861	1 AU140971	AU140971 AU140971
22	727.4	33.9	737	1 AU140993	AU140993 AU140993

23	723.2	33.7	731	1	AU140526	AU140526
24	718.2	33.5	739	1	CA706215	CA706215 DKFZp686J
C 25	707.2	32.9	747	6	CA423317	UI-H-PE1-
26	700.2	32.6	729	1	AU140834	AU140834
C 27	693	32.3	695	7	CN419477	CN419477 170004247
28	693	32.3	706	3	BM674187	BM674187 UI-E-E30-
29	686.8	32.0	2107	4	AK054456	Mus muscu
30	686.6	32.0	837	1	AU140973	AU140973
31	680.8	31.7	699	1	AU140889	AU140889
32	676.6	31.5	734	1	AU140802	AU140802
33	676	31.5	676	1	AL603362	DKFZp686C
34	673.2	31.4	705	1	AU140789	AU140789
35	669	31.2	723	7	CN419622	CN419622 170005319
36	664	30.9	667	1	AU140910	AU140910
37	663.2	30.9	911	5	BX398838	BX398838
C 38	662	30.8	738	6	CD613781	55110394J
39	660.6	30.8	670	5	BU620723	UI-H-FL1-
40	656.4	30.6	836	7	CN538822	UI-M-H50-
41	655.4	30.5	1038	5	BX439175	BX439175
42	655.2	30.5	673	1	AU140450	AU140450
43	654	30.5	851	1	AU141024	AU141024
44	652	30.4	859	6	CF616056	AGENCOURT
45	646.4	30.1	796	1	AU140814	AU140814

ALIGNMENTS

RESULT 1
HSM806653
LOCUS HSM806653 7868 bp mRNA linear HTC 20-JAN-2005
DEFINITION Homo sapiens mRNA; cDNA DKFZp686M04163 (from clone DKFZp686M04163).
ACCESSION BX640608
VERSION BX640608.1 GI:34364616
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7868)
AUTHORS Bloeker, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Oeang, A., Fobo, G., Han, M. and Wiemann, S.
CONSRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (20-JAN-2005) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686M04163) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M04163
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES

Location/Qualifiers
1..7868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686M04163Q"
/db_xref="taxon:9606"
/clone="DKFZp686M04163"
/tissue_type="uterus endothel, primary cell culture"
/clone_lib="686 (synonym: hlccc3). Vector pSport1_sfi; host DH10B; gates SfiIA + SfiIB"
/dev_stage="adult"
/note="fibronectin precursor"
1..7868
/gene="DKFZp686M04163"
267..7157
gene
CDS

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 16:02:29 ; Search time 1207 Seconds
(without alignments)
11855.091 Million cell updates/sec

Title: US-09-581-651d-3
Perfect score: 2147
Sequence: 1 caaactgttggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002as.*
7: Geneseq2002bs.*
8: Geneseq2003as.*
9: Geneseq2003bs.*
10: Geneseq2003cs.*
11: Geneseq2003ds.*
12: Geneseq2004as.*
13: Geneseq2004bs.*
14: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2147	100.0	2147	2	AAx81299 Human mig
2	2143.8	99.9	2358	13	Adr67201 Human bla
3	2119.8	98.7	2443	13	Adq38575 Human SNP
4	2119	98.7	2127	10	Adi18477 Human pro
5	2085.6	97.1	2402	14	Adz26738 Human fib
6	2084.8	96.2	2488	13	Adq38581 Human SNP
7	1975.8	92.0	4295	8	Acc72037 BC00770 g
8	1925.8	89.7	1929	13	Adsi17488 Nucleotid
9	1925.8	89.7	1929	13	Adr97657 Human fib
10	1925.8	89.7	1929	14	Adw44479 Human fib
11	1925.8	89.7	1929	14	Ady55703 Human fib
12	1925.8	89.7	1929	14	Ady32457 Human fib
13	1925.8	89.7	1929	14	Aeb56257 Ewent seq
14	1925.8	89.7	1929	14	Aeb78061 Human DNA
15	1890.6	88.1	7242	13	Adr66637 Human pro
16	1890.6	88.1	7242	13	Adr65953 Human pro
17	1890.6	88.1	7550	8	Acc00412 Human cel
18	1890.6	88.1	7912	14	Adz26740 Human fib
19	1890.6	88.1	8027	11	Adp64998 Human fib

20	1890.6	88.1	8027	12	ADG89565	Adg89565 Human fib
21	1890.6	88.1	8027	12	ADQ29601	Adq29601 Human col
22	1890.6	88.1	8027	13	ADR67200	Adr67200 Human bla
23	1890.6	88.1	8027	14	ADV70129	Adv70129 Tumor-ass
24	1890.6	88.1	8027	14	AEA04397	Aea04397 Human cdn
25	1890.6	88.1	8027	14	AEB17605	Aeb17605 Human fib
26	1890.6	88.1	8272	14	ADZ26744	Adz26744 Human fib
27	1890.6	88.1	8374	14	ADZ26748	Adz26748 Human fib
28	1890.6	88.1	8449	14	ADZ26564	Adz26564 Human fib
29	1890.6	88.1	8647	14	ADZ26742	Adz26742 Human fib
30	1890.6	88.1	8815	14	ADW81010	Adw81010 Fibronect
31	1890.6	88.1	8815	14	ADX69342	Adx69342 DNA encod
32	1890.6	88.1	8815	14	ADZ26746	Adz26746 Human fib
33	1889.8	88.0	6510	13	ADQ38578	Adq38578 Human SNP
34	1889.8	88.0	7823	13	ADQ38582	Adq38582 Human SNP
35	1889.8	88.0	7848	13	ADQ38583	Adq38583 Human SNP
36	1889.8	88.0	7935	13	ADQ38577	Adq38577 Human SNP
37	1889.8	88.0	7959	13	ADQ38586	Adq38586 Human SNP
38	1889.8	88.0	8013	13	ADQ38579	Adq38579 Human SNP
39	1889.8	88.0	8155	13	ADQ38584	Adq38584 Human SNP
40	1889.8	88.0	8226	13	ADQ38574	Adq38574 Human SNP
41	1889.8	88.0	8278	13	ADQ38587	Adq38587 Human SNP
42	1889.8	88.0	8332	13	ADQ38576	Adq38576 Human SNP
43	1889.8	88.0	8371	13	ADQ38580	Adq38580 Human SNP
44	1887.4	87.9	7795	10	ADJ56196	Adj56196 Zebrafish
45	1887.4	87.9	7867	4	AAI57803	AAi57803 Human poi

ALIGNMENTS

RESULT 1
AAx81299

ID AAX81299 standard; cDNA; 2147 BP.

XX AC AAX81299;

XX DT 21-SEP-1999 (first entry)

XX DE Human migration stimulating factor (MSF) 1-alpha encoding cDNA.

XX KW Migration stimulatory factor; MSF; cell migration; modulation; human;

XX KW wound healing; scarring; MSFI-alpha; ss.

XX OS Homo sapiens.

XX PN WO9931233-A1.

XX PD 24-JUN-1999.

XX PF 15-DEC-1998; 98WO-GB003766.

XX PR 16-DEC-1997; 97GB-00026539.

XX PA (UYDU-) UNIV DUNDEE.

XX PI Schor SL, Schor AM;

XX DR WPI; 1999-430039/36.

XX DR P-PSDB; AAY28901.

XX PT Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX PS Example 1; Fig 1; 86pp; English.

XX CC The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. The present sequence represents the nucleotide sequence encoding a human MSFI-alpha protein

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 1, 2005, 12:35:07 ; Search time 48 Seconds
(without alignments)
1105.787 Million cell updates/sec

Title: US-09-581-651D-2
Perfect score: 3687
Sequence: 1 MLRGPGLLLAVQCLGTA.....ISKYLWRPVSIPPRNLGY 642

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3595.5	97.5	2386	2	US-09-961-403-1
2	3579.5	97.1	2386	1	US-09-016-366A-12
3	3578.5	97.1	2446	1	US-08-551-356-2
4	3578.5	97.1	2446	4	PCT-US93-12687-2
5	3439.5	93.3	2355	2	US-10-360-101-235
6	3433.5	93.1	2327	6	5455158-1
7	3422.5	92.8	2324	1	US-08-283-857-1
8	3422.5	92.8	2324	4	PCT-US95-09819-1
9	3406.5	92.4	2231	1	US-08-153-799-16
10	371	10.1	65	1	US-08-982-597A-19
11	371	10.1	65	2	US-09-136-218-19
12	298	8.1	663	2	US-09-194-468A-30
13	297	8.1	631	2	US-08-448-489-17
14	297	8.1	631	2	US-09-689-730-17
15	297	8.1	660	2	US-08-704-711A-18
16	297	8.1	660	2	US-09-521-220-18
17	297	8.1	660	2	US-09-391-104-19
18	297	8.1	660	2	US-09-917-254-89
19	297	8.1	660	2	US-09-949-016-6512
20	297	8.1	660	2	US-09-949-016-7937
21	297	8.1	660	2	US-10-153-185-14
22	295	8.0	422	1	US-08-836-854-12
23	294.5	8.0	188	1	US-08-142-449B-14
24	294.5	8.0	474	1	US-08-836-854-9
25	292.5	7.9	429	2	US-09-194-468A-45
26	290	7.9	48	1	US-08-582-597A-22
27	290	7.9	48	2	US-09-136-218-22

28	285.5	7.7	707	2	US-08-704-711A-19	Sequence 19, Appl
29	285.5	7.7	707	2	US-09-521-220-19	Sequence 19, Appl
30	285.5	7.7	707	2	US-09-391-104-20	Sequence 20, Appl
31	285.5	7.7	707	2	US-09-949-016-6575	Sequence 6575, Ap
32	285.5	7.7	708	2	US-08-448-489-16	Sequence 16, Appl
33	285.5	7.7	708	2	US-09-689-730-16	Sequence 16, Appl
34	285.5	7.7	713	2	US-09-949-016-10629	Sequence 10629, A
35	278	7.5	48	1	US-08-982-597A-20	Sequence 20, Appl
36	278	7.5	48	2	US-09-136-218-20	Sequence 20, Appl
37	258	7.0	43	1	US-08-982-597A-21	Sequence 21, Appl
38	258	7.0	43	2	US-09-136-218-21	Sequence 21, Appl
39	220	6.0	96	1	US-08-717-169-2	Sequence 2, Appl
40	220	6.0	96	2	US-09-228-901A-2	Sequence 2, Appl
41	219.5	6.0	60	1	US-08-982-597A-18	Sequence 18, Appl
42	219.5	6.0	60	2	US-09-136-218-18	Sequence 18, Appl
43	208	5.6	1455	2	US-08-840-062-5	Sequence 5, Appl
44	206	5.6	37	1	US-08-982-597A-17	Sequence 17, Appl
45	206	5.6	37	2	US-09-136-218-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961.403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1

Query Match	97.5%	Score	3595.5	DB 2	Length	2386	
Best Local Similarity	97.2%	Pred. No.	2.6e-301				
Matches	629	Conservative	1	Mismatches	2	Indels	15
Gaps							1
Qy	1	MLRGPGLLLLLAVQCLGTA	VPSTGASKSRQAQMVQPSQSPVAVSQSKPCGYDNGKHQ	60			
Db	1	MLRGPGLLLLLAVQCLGTA	VPSTGASKSRQAQMVQPSQSPVAVSQSKPCGYDNGKHQ	60			
Qy	61	INQWERTYLGNALVCTCYGSGRGNFCESKPEAETCFDKYTGNTYRVGDTYERPKDSMI	120				
Db	61	INQWERTYLGNALVCTCYGSGRGNFCESKPEAETCFDKYTGNTYRVGDTYERPKDSMI	120				
Qy	121	WDCTCIAGRGRICTIANRCHEGGQSYKIGDTRWRPHETGGYMLCVCCLNGKGWTC	180				
Db	121	WDCTCIAGRGRICTIANRCHEGGQSYKIGDTRWRPHETGGYMLCVCCLNGKGWTC	180				
Qy	181	PIAKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY	240				
Db	181	PIAKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY	240				
Qy	241	RIGDTWSKDNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPFTDVAAYVQOPHP	300				
Db	241	RIGDTWSKDNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPFTDVAAYVQOPHP	300				
Qy	301	QPPPYGHCVTDSDGVVYSGWMLKTOGNKMLCTCLNGVYSCQETATVQTYGGNSNGEPC	360				

Db 301 PPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAATVQTYYGNSNGEPC 360
Qy 361 VLPFTYNDRT-----DSTTSNYEODQKYSFCTDHTVLVQTRGNSNGALC 405
Db 361 VLPFTYNGRTFYSCCTEGRQDGLWCSTTSNYEODQKYSFCTDHTVLVQTRGNSNGALC 420
Qy 406 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 465
Db 421 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 480
Qy 466 GDQWDKQDMDGMMRCTCVNGRGSWTCTIAYSQLRQDQICVDDITYNVNDTFFKRHEEGHM 525
Db 481 GDQWDKQDMDGMMRCTCVNGRGSWTCTIAYSQLRQDQICVDDITYNVNDTFFKRHEEGHM 540
Qy 526 LNCTCFGQGRGWKCDPVDQCDSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 585
Db 541 LNCTCFGQGRGWKCDPVDQCDSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 600
Qy 586 PLQTPSSSGPVEVFTITPSPNSHPIQWNAPOPSHISKYLRLWRP 632
Db 601 PLQTPSSSGPVEVFTITPSPNSHPIQWNAPOPSHISKYLRLWRP 647

RESULT 2

US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431

GENERAL INFORMATION:

; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:

INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-016-366A-12

Query Match 97.1%; Score 3579.5; DB 1; Length 2386;
Best Local Similarity 96.8%; Pred. NO. 6.3e-300;
Matches 626; Conservative 1; Mismatches 5; Indels 15; Gaps 1;

Qy 1 MLRGGFGLLLAVQCLGTAVPSTGASKSKRQAMQVQPOS PVAVSQSKPCYDNGKHQY 60
Db 1 MLRGGFGLLLAVLCLGTAVPSTGASKSKRQAMQVQPOS PVAVSQSKPCYDNGKHQY 60
Qy 61 INQWERTYLGNAUVCTCYGSGRGNCEKPEABETCFDKYTGNTYRVGDTYERPKDSMI 120
Db 61 INQWERTYLGNAUVCTCYGSGRGNCEKPEABETCFDKYTGNTYRVGDTYERPKDSMI 120
Qy 121 WDCTCICAGGRISCTTANRCHGEGQSYKIGDTRRRPHETGGYMLECVCLGNGKGEWTC 180
Db 121 WDCTCICAGGRISCTTANRCHGEGQSYKIGDTRRRPHETGGYMLECVCLGNGKGEWTC 180
Qy 181 PIAKCFDHAAGTSYVVGETWEKPYQGMWVDCCTCLGEGSGRICTSRNRCNDQDTRTSY 240
Db 181 PIAKCFDHAAGTSYVVGETWEKPYQGMWVDCCTCLGEGSGRICTSRNRCNDQDTRTSY 240
Qy 241 RIGDTSKKNRGNLLQICITGNGRGWKERHTSVQTTSSGSGPFTDRAAAYQPOPH 300
Db 241 RIGDTSKKNRGNLLQICITGNGRGWKERHTSVQTTSSGSGPFTDRAAAYQPOPH 300
Qy 301 PPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAATVQTYYGNSNGEPC 360
Db 301 PPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAATVQTYYGNSNGEPC 360
Qy 361 VLPFTYNDRT-----DSTTSNYEODQKYSFCTDHTVLVQTRGNSNGALC 405
Db 361 VLPFTYNGRTFYSCCTEGRQDGLWCSTTSNYEODQKYSFCTDHTVLVQTRGNSNGALC 420
Qy 406 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 465
Db 421 HFPFLYNNHNTYDCTSEGRDNMKWCGTTQNYDADQKFGFCFMAAHEEICTTNEGVMYRI 480
Qy 466 GDQWDKQDMDGMMRCTCVNGRGSWTCTIAYSQLRQDQICVDDITYNVNDTFFKRHEEGHM 525
Db 481 GDQWDKQDMDGMMRCTCVNGRGSWTCTIAYSQLRQDQICVDDITYNVNDTFFKRHEEGHM 540
Qy 526 LNCTCFGQGRGWKCDPVDQCDSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 585
Db 541 LNCTCFGQGRGWKCDPVDQCDSTGTFFYQIGDSWEKYVHGVRYQCYCYGRGIGEWHCQ 600
Qy 586 PLQTPSSSGPVEVFTITPSPNSHPIQWNAPOPSHISKYLRLWRP 632
Db 601 PLQTPSSSGPVEVFTITPSPNSHPIQWNAPOPSHISKYLRLWRP 647

RESULT 3

US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700

GENERAL INFORMATION:

; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2005, 18:06:49 ; Search time 405 Seconds
(without alignments)
9423.271 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaactggggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUS_COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1887.4	87.9	8044	3	US-09-566-921-135
2	1826.8	85.1	7803	2	US-08-551-356-1
3	1826.8	85.1	7803	6	PCT-US93-12687-1
4	1743.4	81.2	7679	3	US-09-220-132-38
5	1743.4	81.2	7680	3	US-09-023-655-1289
6	1743.4	81.2	7680	6	PCT-US93-09819-6
7	1740.2	81.1	7705	2	US-08-259-569-16
8	1740.2	81.1	7705	2	US-08-826-885-16
9	1738.6	81.0	7705	9	5455158-2
10	671.2	31.3	986	2	US-07-637-250A-8
11	671.2	31.3	986	2	US-08-145-061-8
12	97.4	4.5	186	2	US-08-153-799-5
13	73.8	3.4	2109	3	US-09-799-451-345
14	73.8	3.4	2334	3	US-09-023-655-996
15	73.8	3.4	2334	3	US-09-949-016-704
16	73.8	3.4	2335	3	US-09-799-451-346
17	73.8	3.4	2335	3	US-09-949-016-4758
18	73.2	3.4	11665	3	US-09-949-016-12446
19	73.2	3.4	11665	3	US-09-949-016-16500
20	72.4	3.4	85	2	US-08-259-569-28
21	72.4	3.4	85	2	US-08-826-885-18
22	72.2	3.4	2334	2	US-08-457-304A-33
23	72.2	3.4	2334	2	US-08-456-701A-33
24	72.2	3.4	2334	3	US-08-684-932A-33

```

25 72 3.4 72 2 US-08-259-569-29 Sequence 29, Appl
26 72 3.4 72 2 US-08-826-885-29 Sequence 29, Appl
27 72 3.4 72 2 5455158-9 Patent No. 5455158
28 71.8 3.3 1983 3 US-09-949-016-2066 Sequence 2066, Ap
29 71.8 3.3 3069 3 US-09-949-016-641 Sequence 641, App
30 69 3.2 69 2 US-08-259-569-23 Sequence 23, Appl
31 69 3.2 69 2 US-08-826-885-23 Sequence 23, Appl
32 65.8 3.1 69 9 5455158-6 Patent No. 5455158
33 65.4 3.0 69 2 US-08-259-569-30 Sequence 30, Appl
34 65.4 3.0 69 2 US-08-826-885-30 Sequence 30, Appl
35 65.2 3.0 601 3 US-09-949-016-30391 Sequence 30391, A
36 65.2 3.0 29954 3 US-09-949-016-70165 Sequence 70165, A
37 65.2 3.0 29954 3 US-09-949-016-13808 Sequence 13808, A
38 65.2 3.0 31040 3 US-09-949-016-12383 Sequence 12383, A
39 65 3.0 75 2 US-08-259-569-18 Sequence 18, Appl
40 65 3.0 75 2 US-08-826-885-18 Patent No. 5455158
41 64.8 3.0 68 9 5455158-4 Sequence 27, Appl
42 61.8 2.9 71 2 US-08-259-569-27 Sequence 27, Appl
43 61.8 2.9 71 2 US-08-826-885-27 Patent No. 5455158
44 61.8 2.9 71 9 5455158-8 Sequence 17, Appl
45 61.2 2.9 71 2 US-08-259-569-17

```

ALIGNMENTS

```

RESULT 1
US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566,921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

```

```

Query Match 87.9%; Score 1887.4; DB 3; Length 8044;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 6; Indels 45; Gaps 1;
Qy 1 CAAACTTGGTGGCAACTTGGCTCCCGTGGGGGGTCTCTCCCGCCAGGTCCTCAACATGC 60
Db 213 CAAACTTGGTGGCAACTTGGCTCCCGTGGGGGGTCTCTCCCGCCAGGTCCTCAACATGC 272
Qy 61 TTAGGGGTCCGGGGCCCGGGCTGCTGCTGCTGGCCGCTCCAGTGCCTGGGGACACGGGTGC 120
Db 273 TTAGGGGTCCGGGGCCCGGGCTGCTGCTGCTGGCCGCTCTGCTGCTGGGACACGGGTGC 332
Qy 121 CCTCCACGGAGGCTCGAAGCAAGAGCAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCC 180
Db 333 CCTCCACGGAGGCTCGAAGCAAGAGCAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCC 392
Qy 181 CGTGGGTGTCAGTCAAAAGCAAGCCCGTTGTTATGCAATGGAACACTATCAGATAA 240
Db 393 CGTGGGTGTCAGTCAAAAGCAAGCCCGTTGTTATGCAATGGAACACTATCAGATAA 452
Qy 241 ATCAACAGTGGGAGCGGACCTACTAGGCAATGGTTGGTTGTACTTGTATGGAGAA 300
Db 453 ATCAACAGTGGGAGCGGACCTACTAGGCAATGGTTGGTTGTACTTGTATGGAGAA 512

```

QY 301 GCCAGGTTTAACTCGAGAGTAACCTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACA 360
DB 513 GCCAGGTTTAACTCGAGAGTAACCTGAAGCTGAAGAGACTTGTCTTTGACAAAGTACA 572
QY 361 CTGGGAAACATTTACCGAGTGGGTGACATTTATGAGCGTCTTAAAGACTCCATGATCTGGG 420
DB 573 CTGGGAAACATTTACCGAGTGGGTGACATTTATGAGCGTCTTAAAGACTCCATGATCTGGG 632
QY 421 ACTGTACCTGCATCGGGCTGGCGAGGGAGAAATGAAGCTGTACCATCCAAACCGCTGCC 480
DB 633 ACTGTACCTGCATCGGGCTGGCGAGGGAGAAATGAAGCTGTACCATCCAAACCGCTGCC 692
QY 481 ATGAAGGGGGTCACTTCCTACAAGATTGGTGACACCTCGAGGAGACCAATGAGACTGGTG 540
DB 693 ATGAAGGGGGTCACTTCCTACAAGATTGGTGACACCTCGAGGAGACCAATGAGACTGGTG 752
QY 541 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGAGGAAATGGAACCTGCAAGCCCA 600
DB 753 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGAGGAAATGGAACCTGCAAGCCCA 812
QY 601 TAGCTGAGAGTGTGTGTCTTGCATCATGCTGCTGGGACTTCTATGTGGTTCGAGAAAGCTGGG 660
DB 813 TAGCTGAGAGTGTGTGTCTTGCATCATGCTGCTGGGACTTCTATGTGGTTCGAGAAAGCTGGG 872
QY 661 AGAAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTGCTGGGAGAGCGACGGAC 720
DB 873 AGAAGCCCTACCAAGGCTGGATGATGTTGATGTTGCTGCTGGGAGAGCGACGGAC 932
QY 721 GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACCAAGGACATCTTATAGAA 780
DB 933 GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACCAAGGACATCTTATAGAA 992
QY 781 TTGGAGACACCTGGAGAGAGGAGTAATCAGGAAAGCTGCTCAGTGCACTTGCACAG 840
DB 993 TTGGAGACACCTGGAGAGAGGAGTAATCAGGAAAGCTGCTCAGTGCACTTGCACAG 1052
QY 841 GCAACGGCGGAGAGAGTGGAGTGTGAGAGGACACACCTCTGTGAGGACCAATCGAGCG 900
DB 1053 GCAACGGCGGAGAGAGTGGAGTGTGAGAGGACACACCTCTGTGAGGACCAATCGAGCG 1112
QY 901 GATCTGGCCCTTCAACCGATGTTCTGAGCTGTTTACCAACCGAGCTTACCCCCAGC 960
DB 1113 GATCTGGCCCTTCAACCGATGTTCTGAGCTGTTTACCAACCGAGCTTACCCCCAGC 1172
QY 961 CTCCTCCCTATGGCCACTGTGTACAGACAGTGTGTCTCTCTGTGGGATGCACT 1020
DB 1173 CTCCTCCCTATGGCCACTGTGTACAGACAGTGTGTCTCTCTGTGGGATGCACT 1232
QY 1021 GGCTGAAGACACAAAGGAAATAGCAAAATGCTTTGACAGCTGCTGGGCAACGGAGTCAGT 1080
DB 1233 GGCTGAAGACACAAAGGAAATAGCAAAATGCTTTGACAGCTGCTGGGCAACGGAGTCAGT 1292
QY 1081 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGTCT 1140
DB 1293 GCCAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGTCT 1352
QY 1141 TACCATTTCACTTACACGACGAGC----- 1165
DB 1353 TACCATTTCACTTACAAATGGCAGGAGTCTCTCTGTGACCAACAGAGGGCGACAGGACG 1412
QY 1166 -----GGACAGCAACACTTGAATTTATGAGCAGGACAGCAAAATACTCTTTCTGCA 1215
DB 1413 GACATCTTTGGTGGCAGCAACACTTGAATTTATGAGCAGGACAGCAAAATACTCTTTCTGCA 1472
QY 1216 CAGACACACTTTTGGTTCAGACTCGAGAGGAGAAATTCAAATGGTGTCTTGTGCCACT 1275
DB 1473 CAGACACACTTTTGGTTCAGACTCGAGAGGAGAAATTCAAATGGTGTCTTGTGCCACT 1532
QY 1276 TCCCTCTCTATACACACCAACCAATTACACTGATTGCACTTCTGAGGGCAGAGAGACA 1335
DB 1533 TCCCTCTCTATACACACCAACCAATTACACTGATTGCACTTCTGAGGGCAGAGAGACA 1592
QY 1336 ACATGAAGTGGTGGGACCAACACAGAACTATGATGCCACCAAGATTTGGGTTCTGCC 1395

DB 1593 ACATGAAGTGGTGGGACCAACAGAACTATGATGCCACAGAAATTTGGGTTCTGCC 1652
QY 1396 CCATGGCTGCCACGAGGAAATCTGCACCAACCAATGAAGGGTCTATGTCGCAATGGAG 1455
DB 1653 CCATGGCTGCCACGAGGAAATCTGCACCAACCAATGAAGGGTCTATGTCGCAATGGAG 1712
QY 1456 ATCAGTGGGATAAGCAGCAGCATGGGTCACTGATGAGGTGCACTGTGTGTGGGAATG 1515
DB 1713 ATCAGTGGGATAAGCAGCAGCATGGGTCACTGATGAGGTGCACTGTGTGTGGGAATG 1772
QY 1516 GTCTGGGGAATGACATGCAATGCCCTACTCGAGCTTCGAGATCACTGCAATGTTGTGATG 1575
DB 1773 GTCTGGGGAATGACATGCAATGCCCTACTCGAGCTTCGAGATCACTGCAATGTTGTGATG 1832
QY 1576 ACATCACTTCAATGTCAGCAGCATTCACCAAGCGTTCATGAAGAGGGGACATGCTGA 1635
DB 1833 ACATCACTTCAATGTCAGCAGCATTCACCAAGCGTTCATGAAGAGGGGACATGCTGA 1892
QY 1636 ACTGTACATGCTTCGGTCAAGGTCGGGCGAGGTGGAAGTGTGATCCCGTCGACCAATGCC 1695
DB 1893 ACTGTACATGCTTCGGTCAAGGTCGGGCGAGGTGGAAGTGTGATCCCGTCGACCAATGCC 1952
QY 1696 AGGATTCAAGACTGGGACGTTTTATCAAAATGAGATTCATGGGAGAGTATGTGCAATG 1755
DB 1953 AGGATTCAAGACTGGGACGTTTTATCAAAATGAGATTCATGGGAGAGTATGTGCAATG 2012
QY 1756 GTGTCAAGTACCACTGCTACTGCTATGCGCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 1815
DB 2013 GTGTCAAGTACCACTGCTACTGCTATGCGCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 2072
QY 1816 TACAGACTTCAAGCTCAAGTGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTC 1875
DB 2073 TACAGACTTCAAGCTCAAGTGTCTGTGCGAAGTATTTATCACTGAGACTCCGAGTC 2132
QY 1876 AGCCCAACTCCACCCCATCCAGTGGAAATGACCAAGCCATCTCAANTTCCAGATACA 1935
DB 2133 AGCCCAACTCCACCCCATCCAGTGGAAATGACCAAGCCATCTCAANTTCCAGATACA 2192
QY 1936 TTCTCAGTGGAGACCT 1952
DB 2193 TTCTCAGTGGAGACCT 2209

RESULT 2
US-08-551-356-1
; Sequence 1, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Iran, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/551,356
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE: US/07/998,271
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:49:16 ; Search time 664.314 Seconds
(without alignments)
1625.652 Million cell updates/sec

Title: US-09-581-651D-1
Perfect score: 3781
Sequence: 1 NLVATCLPVRSLPHRLNML.....ISKYILWRPVSIPRNLGY 660

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool.h/US09581651/runat.07112005.092224.28801/app.query.fasta_1.1834
-DB=Issued_Patents_NA -QFAST=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORES=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09581651@cgn 1 187 @runat.07112005.092224.28801 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3686.5	97.5	8044	4	US-09-566-921-135
2	3594.5	94.8	7803	2	Sequence 135, App
3	3594.5	94.8	7803	5	Sequence 1, Appli
4	3448.5	91.2	7679	4	Sequence 38, Appl
5	3448.5	91.2	7680	4	Sequence 1289, Ap
6	3448.5	91.2	7680	5	Sequence 6, Appli
7	3433.5	90.8	7705	1	Sequence 16, Appl
8	3433.5	90.8	7705	2	Sequence 16, Appl
9	3428.5	90.7	7705	6	Patent No. 5455158
10	3428.5	90.7	7705	6	Patent No. 5455158
11	1665.5	44.0	986	1	US-07-637-250A-8
12	1665.5	44.0	986	1	US-08-145-061-8

13	325.5	8.6	2109	4	US-09-799-451-345	Sequence 345, App
14	311	8.2	2335	1	US-09-799-451-346	Sequence 346, App
15	306.5	8.1	2333	1	US-08-392-678-33	Sequence 33, Appl
16	305.5	8.1	2334	1	US-08-457-304A-33	Sequence 33, Appl
17	305.5	8.1	2334	1	US-08-456-701A-33	Sequence 33, Appl
18	305.5	8.1	2334	3	US-08-684-932A-33	Sequence 33, Appl
19	305.5	8.1	2334	4	US-09-023-655-996	Sequence 996, App
20	305.5	8.1	2334	4	US-09-949-016-704	Sequence 704, App
21	305.5	8.1	2335	4	US-09-949-016-4758	Sequence 4758, Ap
22	302	8.0	1983	4	US-09-949-016-2066	Sequence 2066, Ap
23	302	8.0	3069	4	US-09-949-016-641	Sequence 641, App
24	298	7.9	2123	4	US-08-194-468A-29	Sequence 29, Appl
25	294.5	7.8	567	1	US-08-142-449B-5	Sequence 5, Appli
26	225	6.0	11665	4	US-09-949-016-12446	Sequence 12446, A
27	225	6.0	11665	4	US-09-949-016-16500	Sequence 16500, A
28	217.5	5.8	4588	3	US-08-840-062-1	Sequence 1, Appli
29	208	5.5	4627	4	US-09-949-016-935	Sequence 935, App
30	208	5.5	5633	4	US-09-023-655-1490	Sequence 1490, Ap
31	199.5	5.3	4852	4	US-09-976-594-167	Sequence 167, App
32	199.5	5.3	5158	4	US-09-023-655-1347	Sequence 1347, Ap
33	197	5.2	4621	4	US-09-949-016-4577	Sequence 4577, Ap
34	197	5.2	5627	4	US-09-949-016-2444	Sequence 2444, Ap
35	196.5	5.2	29954	4	US-09-949-016-13808	Sequence 13808, A
36	196.5	5.2	31040	4	US-09-949-016-12383	Sequence 12383, A
37	193.5	5.1	4464	2	US-08-400-159-7	Sequence 7, Appli
38	193.5	5.1	4483	3	US-08-611-729A-7	Sequence 7, Appli
39	193.5	5.1	4483	4	US-09-195-524-7	Sequence 7, Appli
40	193	5.1	3955	3	US-09-214-278-4	Sequence 4, Appli
41	193	5.1	3955	4	US-09-855-722-4	Sequence 4, Appli
42	189.5	5.0	4771	3	US-08-840-062-3	Sequence 3, Appli
43	189	5.0	601	4	US-09-949-016-30388	Sequence 30388, A
44	189	5.0	601	4	US-09-949-016-30391	Sequence 30391, A
45	189	5.0	601	4	US-09-949-016-70162	Sequence 70162, A

ALIGNMENTS

RESULT 1

US-09-566-921-135
; Sequence 135, Application US/09566921
; Patent No. 6682888
; GENERAL INFORMATION:
; APPLICANT: Loring, Jeanne F.
; APPLICANT: Tingley, Debora W.
; APPLICANT: Edwards, Carla M.
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
; FILE REFERENCE: PA-0024 US
; CURRENT APPLICATION NUMBER: US/09/566, 921
; CURRENT FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 135
; LENGTH: 8044
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14
US-09-566-921-135

Alignment Scores:
Pred. No.: 0
Score: 3686.50
Percent Similarity: 97.29%
Best Local Similarity: 97.29%
Query Match: 97.50%
DB: 4
Length: 8044
Matches: 647
Conservative: 0
Mismatches: 3
Indels: 15
Gaps: 1

US-09-581-651D-1 (1-660) x US-09-566-921-135 (1-8044)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAenMetLeu 20
Db 215 AACTTGTGGCACTTGCCTCCCGGCGGGCTCTCTCCCCACCCTCTCAACATGCTT 274

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2005, 18:42:04 ; Search time 1153 Seconds
(without alignments)
11023.144 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaactgttggaacttcg.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2147	100.0	2147	2	AX81299 Human mig
2	2143.8	99.9	2358	13	Adr67201 Human bla
3	2119.8	98.7	2443	13	Adq38575 Human SNP
4	2119	98.7	2127	10	Adi18477 Human pro
5	2064.8	96.2	2488	13	Adq38581 Human SNP
6	1975.8	92.0	4295	8	ACC72037 BCU0770 g
7	1925.8	89.7	1929	13	Adsi17488 Nucleotid
8	1925.8	89.7	1929	13	Adr97657 Human fib
9	1890.6	88.1	7242	13	Adr66637 Human pro
10	1890.6	88.1	7242	13	Adr65953 Human pro
11	1890.6	88.1	7550	8	ACC00412 Human cel
12	1890.6	88.1	8027	11	Adp64998 Human fib
13	1890.6	88.1	8027	12	ADG89565 Human fib
14	1890.6	88.1	8027	13	Adq29601 Human col
15	1890.6	88.1	8027	13	Adr67200 Human bla
16	1889.8	88.0	6510	13	Adq38578 Human SNP
17	1889.8	88.0	7823	13	Adq38582 Human SNP
18	1889.8	88.0	7848	13	Adq38583 Human SNP
19	1889.8	88.0	7935	13	Adq38577 Human SNP
20	1889.8	88.0	7959	13	Adq38586 Human SNP

21	1889.8	88.0	8013	13	ADQ38579	Adq38579 Human SNP	
22	1889.8	88.0	8155	13	ADQ38584	Adq38584 Human SNP	
23	1889.8	88.0	8226	13	ADQ38574	Adq38574 Human SNP	
24	1889.8	88.0	8278	13	ADQ38587	Adq38587 Human SNP	
25	1889.8	88.0	8332	13	ADQ38576	Adq38576 Human SNP	
26	1889.8	88.0	8371	13	ADQ38580	Adq38580 Human SNP	
27	1887.4	87.9	7795	10	ADJ56196	Adj56196 Zebrafish	
28	1887.4	87.9	7867	4	AAI57803	AAI57803 Human pol	
29	1887.4	87.9	8044	6	RA594866	RA594866 Human DNA	
30	1887.4	87.9	8044	12	ADI61767	ADI61767 Human cDN	
31	1887.4	87.9	8062	4	AAI57802	AAI57802 Human pol	
32	1887.4	87.9	8137	4	AAI57805	AAI57805 Human pol	
33	1887.4	87.9	8230	4	AAI57804	AAI57804 Human pol	
C	34	1858.6	86.6	8066	4	AAI59591	AAI59591 Human pol
C	35	1858.6	86.6	8066	4	AAI59589	AAI59589 Human pol
C	36	1858.6	86.6	8066	4	AAI59588	AAI59588 Human pol
C	37	1858.6	86.6	8066	4	AAI59590	AAI59590 Human pol
38	1839.6	85.7	7361	9	ACD06170	ACD06170 Human cDN	
39	1838.8	85.6	8039	3	AA898889	AA898889 Human FN	
40	1826.8	85.1	7803	2	AAQ70009	AAQ70009 Fibrinoge	
41	1814.6	84.5	8216	5	RA586466	RA586466 DNA encod	
42	1754.6	81.7	7677	13	ACN37820	ACn37820 Tumour-as	
43	1743.4	81.2	4860	3	AAA35009	AAa35009 Human ade	
44	1743.4	81.2	6988	9	ACD06169	ACD06169 Human cDN	
45	1743.4	81.2	7020	13	ADR90518	AdR90518 Human cDN	

ALIGNMENTS

RESULT 1
AX81299
ID AX81299 standard; CDNA; 2147 BP.
XX
AC AX81299;
XX
DT 21-SEP-1999 (first entry)
XX
DE Human migration stimulating factor (MSF) 1-alpha encoding cDNA.
XX
KW Migration stimulating factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; 88.
XX
OS Homo sapiens.
XX
PN WO9931233-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-GB003766.
XX
PR 16-DEC-1997; 97GB-00026539.
XX
PA (UYDU-) UNIV DUNDEE.
XX
FI Schor SL, Schor AM;
XX
DR WPI; 1999-430039/36.
DR P-PSDB; RAY28901.
XX
PT Proteins with cell migration stimulatory activity used in treating wound
PT and preventing scarring.
XX
PS Example 1; Fig 1; 86pp; English.
XX
CC The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. The present sequence
CC represents the nucleotide sequence encoding a human MSF1-alpha protein
XX
SQ Sequence 2147 BP; 575 A; 512 C; 575 G; 485 T; 0 U; 0 Other;

Query Match									
Best Local Similarity 100.0%; Score 2147; DB 2; Length 2147;									
Matches 2147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CAAACTTGGTGCAACTTGGCTCCCGGTGCGGGCTCTCTCCCCACCGTCTCAACATGC	60						
Db	1	CANACTTGGTGCAACTTGGCTCCCGGTGCGGGCTCTCTCCCCACCGTCTCAACATGC	60						
Qy	61	TTAGGGTCCGGGGCCCGGGCTGTGCTGTGGCCGTCAAGTCTGGGGACAGCGGTGC	120						
Db	61	TTAGGGTCCGGGGCCCGGGCTGTGCTGTGGCCGTCAAGTCTGGGGACAGCGGTGC	120						
Qy	121	CCTCAACGGGAGCTCGAAGACAGAGCGCAGGCTCAGCAATGTTTTCAGCCCAAGTCCC	180						
Db	121	CCTCAACGGGAGCTCGAAGACAGAGCGCAGGCTCAGCAATGTTTTCAGCCCAAGTCCC	180						
Qy	181	CGGTGGCTGTCACTCAAAAGCAAGCCCGTGTGTATGACAAATGGAAAAACATATCAGATAA	240						
Db	181	CGGTGGCTGTCACTCAAAAGCAAGCCCGTGTGTATGACAAATGGAAAAACATATCAGATAA	240						
Qy	241	ATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTGGTGTGTTTGTACTTGTATGAGGAA	300						
Db	241	ATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTGGTGTGTTTGTACTTGTATGAGGAA	300						
Qy	301	CCCGAGGTTTTAACTGCGAGAGTAAACCTGAGCTGGAAGACTTGTGACAAAGTACA	360						
Db	301	CCCGAGGTTTTAACTGCGAGAGTAAACCTGAGCTGGAAGACTTGTGACAAAGTACA	360						
Qy	361	CTGGGAAACATTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCGGG	420						
Db	361	CTGGGAAACATTTACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCGGG	420						
Qy	421	ACTGTACTGTGCATCGGGCTGGGCGAGGGAGAAATAGCTGTACCATCGCAACCCGTGCC	480						
Db	421	ACTGTACTGTGCATCGGGCTGGGCGAGGGAGAAATAGCTGTACCATCGCAACCCGTGCC	480						
Qy	481	ATGAAGGGGTTCAGTCTCAAGAATGGTGAACCTTGGAGGAGACCAATGAGATCTGGT	540						
Db	481	ATGAAGGGGTTCAGTCTCAAGAATGGTGAACCTTGGAGGAGACCAATGAGATCTGGT	540						
Qy	541	GTTACATGTTAGAGTGTGTCTTTGGTAAATGGAAAAAGGAATGGACCTGCAAGCCCA	600						
Db	541	GTTACATGTTAGAGTGTGTCTTTGGTAAATGGAAAAAGGAATGGACCTGCAAGCCCA	600						
Qy	601	TAGCTGAGAAAGTGTGTTGATCATGCTGTGGGACTTCTTATGTGGTGGGAGAAAAGTGGG	660						
Db	601	TAGCTGAGAAAGTGTGTTGATCATGCTGTGGGACTTCTTATGTGGTGGGAGAAAAGTGGG	660						
Qy	661	AGAAAGCCCTACCAAGGCTGGATGGTATGTTGTTGCTGGGAGGAGGACGGGAC	720						
Db	661	AGAAAGCCCTACCAAGGCTGGATGGTATGTTGTTGCTGGGAGGAGGACGGGAC	720						
Qy	721	GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACCAAGGACATCTATAGAA	780						
Db	721	GCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACCAAGGACATCTATAGAA	780						
Qy	781	TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAAACCTGCTCCAGTGCATCTGCACAG	840						
Db	781	TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAAACCTGCTCCAGTGCATCTGCACAG	840						
Qy	841	GCAACGGCCGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGCAAGACCAATCGAGCG	900						
Db	841	GCAACGGCCGAGAGTGGAAAGTGTGAGAGGACACCTCTGTGCAAGACCAATCGAGCG	900						
Qy	901	GATCTGGCCCTTCAACCGATGTTGCTGCACTGTTTACCAACCGCAGCCTCAACCCCGAGC	960						
Db	901	GATCTGGCCCTTCAACCGATGTTGCTGCACTGTTTACCAACCGCAGCCTCAACCCCGAGC	960						
Qy	961	CTCTCCCTCATGGCCACTGTGTACAGACAGTGGTGTGTACTCTGTGGGAGTGCAGT	1020						
Db	961	CTCTCCCTCATGGCCACTGTGTACAGACAGTGGTGTGTACTCTGTGGGAGTGCAGT	1020						

Qy	1021	GGCTGAAGACACAAAGGAAATAAGCAAATGCTTTGCAGTGCCTTGGGCAACGGAGTCAGCT	1080
Db	1021	GGCTGAAGACACAAAGGAAATAAGCAAATGCTTTGCAGTGCCTTGGGCAACGGAGTCAGCT	1080
Qy	1081	GCCAAGAGACAGCTGTAAACCCAGACTTAACCGTGGCAACTCAAAATGGAGAGCATGTGTCT	1140
Db	1081	GCCAAGAGACAGCTGTAAACCCAGACTTAACCGTGGCAACTCAAAATGGAGAGCATGTGTCT	1140
Qy	1141	TACCATTCACCTCAACGACAGGACGAGCAGCAAACTTCGAAATTAAGAGCAGGACCAAGA	1200
Db	1141	TACCATTCACCTCAACGACAGGACGAGCAGCAAACTTCGAAATTAAGAGCAGGACCAAGA	1200
Qy	1201	AATACTCTTCTGCACAGACCACTGCTTTTGGTTCAGACTCGAGGAGGAAATTCCAATG	1260
Db	1201	AATACTCTTCTGCACAGACCACTGCTTTTGGTTCAGACTCGAGGAGGAAATTCCAATG	1260
Qy	1261	GTGCTTGTGCACTTCCCTTCTTATACAAACCACTTACACTGATTTGCACTTCTG	1320
Db	1261	GTGCTTGTGCACTTCCCTTCTTATACAAACCACTTACACTGATTTGCACTTCTG	1320
Qy	1321	AGGCGAGAGAGACAAATGAAAGTGGTGGGACCAACAGAACTATGATGCCGACCAAGA	1380
Db	1321	AGGCGAGAGAGACAAATGAAAGTGGTGGGACCAACAGAACTATGATGCCGACCAAGA	1380
Qy	1381	AGTTTGGGTTCTGCCCATGCTGCCACGAGGAAATCTGCACAACTCAATGAAGGGTCA	1440
Db	1381	AGTTTGGGTTCTGCCCATGCTGCCACGAGGAAATCTGCACAACTCAATGAAGGGTCA	1440
Qy	1441	TGTACCGCATTTGGAGATCAGTGGGATAGCAGATGATGGGTCAATGATGAGGTGCA	1500
Db	1441	TGTACCGCATTTGGAGATCAGTGGGATAGCAGATGATGGGTCAATGATGAGGTGCA	1500
Qy	1501	CGTGTGTTGGGAAATGGTGGGGAATGGACATGCTTGGCTTCTCGCAGCTTTCAGATC	1560
Db	1501	CGTGTGTTGGGAAATGGTGGGGAATGGACATGCTTGGCTTCTCGCAGCTTTCAGATC	1560
Qy	1561	AGTGCAATTTGATGACATCACTTACAAATGTGAACGACACATTCACAAGCGTCAATGAG	1620
Db	1561	AGTGCAATTTGATGACATCACTTACAAATGTGAACGACACATTCACAAGCGTCAATGAG	1620
Qy	1621	AGGGGCAATGCTGNAATGTTGATGCTTGGTTCAGGTCGGGGCAGGTGGAGTGTGATC	1680
Db	1621	AGGGGCAATGCTGNAATGTTGATGCTTGGTTCAGGTCGGGGCAGGTGGAGTGTGATC	1680
Qy	1681	CGGTGCAACCAATGCCAGGATTCAGAGACTGGGACGTTTTTATCAAAATTCGAGATTCATGGG	1740
Db	1681	CGGTGCAACCAATGCCAGGATTCAGAGACTGGGACGTTTTTATCAAAATTCGAGATTCATGGG	1740
Qy	1741	AGAAATATGTGATGCTGATACCAAGTGTCTACTGCTATGGCGTGGCATTTGGGAGT	1800
Db	1741	AGAAATATGTGATGCTGATACCAAGTGTCTACTGCTATGGCGTGGCATTTGGGAGT	1800
Qy	1801	GGCATTCGCAACCTTTACAGACCTATCCAAGTTCAGTGGTCTCTGTCGAAAGTATTTATCA	1860
Db	1801	GGCATTCGCAACCTTTACAGACCTATCCAAGTTCAGTGGTCTCTGTCGAAAGTATTTATCA	1860
Qy	1861	CTGAGACTCCGAGTCAGCCCACTCCACCCCATCCAGTGGAAATGCCACAGACCATCTC	1920
Db	1861	CTGAGACTCCGAGTCAGCCCACTCCACCCCATCCAGTGGAAATGCCACAGACCATCTC	1920
Qy	1921	ACATTTCCAAAGTACATTTCTCAGGTGGAGACCTGTGAGTATCCACCCAGAAAACCTTGGAT	1980
Db	1921	ACATTTCCAAAGTACATTTCTCAGGTGGAGACCTGTGAGTATCCACCCAGAAAACCTTGGAT	1980
Qy	1981	ACTGAGTCTCTTAATCTTATCAATTTCTGATGGTTTTCTTTTCCAGCTTTTGAGCCAA	2040
Db	1981	ACTGAGTCTCTTAATCTTATCAATTTCTGATGGTTTTCTTTTCCAGCTTTTGAGCCAA	2040
Qy	2041	CAACTCTGATTAACATTTCTTATGCTATGCTATGCTTGTAGTGAACAAACAATATG	2100
Db	2041	CAACTCTGATTAACATTTCTTATGCTATGCTATGCTTGTAGTGAACAAACAATATG	2100
Qy	2101	TGGTCAATTTAAATTTGACTTGTAGACTGAAAAAATAAAAAAAAAAAAAA	2147

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 10, 2005, 20:50:35 ; Search time 9269 Seconds
(without alignments)
11223.794 Million cell updates/sec

Title: US-09-581-651D-3
Perfect score: 2147
Sequence: 1 caaacttgggtggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2147	100.0	2147	6	BD137021 Polypepti
2	2147	100.0	2147	6	AX003229 Sequence
3	2147	100.0	2147	9	HA276395 Homo sapi
4	2143.8	99.9	2358	6	C0875358 Sequence
5	2084	97.1	2192	9	HA535086 Homo sapi
6	2083.4	96.1	4321	9	HA535086 Homo sapi
7	1925.8	89.7	1929	6	C0871810 Sequence
8	1925.8	89.7	1929	6	C0871828 Sequence
9	1890.6	88.1	7753	9	AB191261 Homo sapi
10	1890.6	88.1	8027	6	C0833991 Sequence
11	1890.6	88.1	8027	6	C0875357 Sequence
12	1889	88.0	7868	9	HA535086 Homo sapi
13	1889	88.0	7951	9	HA535086 Homo sapi
14	1889	88.0	7951	9	HA535086 Homo sapi
15	1889	88.0	8411	9	HA535086 Homo sapi
16	1887.4	87.9	7502	9	HA535086 Homo sapi
17	1887.4	87.9	8042	9	HA535086 Homo sapi
18	1887.4	87.9	8044	6	AX454662 Sequence
19	1887.4	87.9	8044	6	AX281712 Sequence

20	1885.8	87.8	7544	9	HA535086 Homo sapi
21	1885.8	87.8	8320	9	HA535086 Homo sapi
22	1884.2	87.8	8030	9	HA535086 Homo sapi
23	1884.2	87.8	8035	9	HA535086 Homo sapi
24	1884.2	87.8	8421	9	HA535086 Homo sapi
25	1826.8	85.1	7803	6	AR051657 Sequence
26	1743.4	81.2	7679	6	AR274901 Sequence
27	1743.4	81.2	7680	6	AR380744 Sequence
28	1743.4	81.2	7680	6	AX277596 Sequence
29	1743.4	81.2	7680	6	AX335368 Sequence
30	1743.4	81.2	7680	9	HSFIB1 Human mRNA
31	1740.2	81.1	7705	6	AL14133 Fibronectin
32	1740.2	81.1	7705	6	AR034630 Sequence
33	1740.2	81.1	7705	6	E01162 cDNA encodi
34	1740.2	81.1	7705	6	I70110 Sequence 16
35	1738.6	81.0	7705	6	AR364992 Sequence
36	1547	72.1	6816	6	C0715726 Sequence
37	1461.8	68.1	3059	10	BC051082 Mus muscu
38	1429.8	66.6	8329	6	AX402055 Sequence
39	1429.8	66.6	8329	10	RNFIBRON
40	1263	58.8	7323	11	BV178397 sqm10026
41	1027.8	47.9	1139	9	HSU42404 Human fibro
42	990.6	46.1	2409	6	C0731571 Sequence
43	975.4	45.4	8313	5	BC072841 Xenopus l
44	972.2	45.3	8216	5	XELFRNCT M77820 Xenopus lae
45	940	43.8	7299	9	HA535086 Homo sapi

ALIGNMENTS

RESULT 1
BD137021
LOCUS BD137021 2147 bp DNA linear PAT 18-SEP-2002
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION BD137021
VERSION BD137021.1 GI:23231966
KEYWORDS JP 2002508179-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2147)
AUTHORS Schor,S.L. and Schor,A.M.
TITLE Polypeptides, polynucleotides and uses thereof
JOURNAL Patent: JP 2002508179-A 1 19-MAR-2002;
UNIVERSITY OF DUNDEE
COMMENT OS Homo sapiens (human)
PN JP 2002508179-A/1
PD 19-MAR-2002
PF 15-DEC-1998 JP 2000539133
PR 16-DEC-1997 GB 9726539.1
PI SETH LAWRENCE SCHOR,ANA MARIA SCHOR
PC C12N15/09,A61K38/00,A61F17/02,C07K14/78,C07K16/18,C12N5/10, PC
C12P21/08,
PC C12Q1/68,G01N33/574,C12N15/00,A61K37/02,C12N5/00 CC
Polypeptides, polynucleotides and uses thereof. FH Key
Location/Qualifiers
FT source 1..2147
FT /organism='Homo sapiens (human)'.
Location/Qualifiers
1..2147
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

FEATURES

source
Query Match 100.0%; Score 2147; DB 6; Length 2147;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Qy 1 CAAACTTGGTGGCAACTTGGCTCCCGTGGGGGGTCTCTCCCGACCGTCTCAACATGC 60
|||||

Ds 1 CAAACTTGGTGGCAACTTGCCTCCGGTGGGGCGTCTCTCCCCCAACGGTCTCAACATGC 60
Qy 61 TTAGGGGTCCGGGGCCCGGGCTGTCTGTCTGGCGGTCCAGTGCCTGGGGACAGCGGTGC 120
Ds 61 TTAGGGGTCCGGGGCCCGGGCTGTCTGTCTGGCGGTCCAGTGCCTGGGGACAGCGGTGC 120
Qy 121 CCTCCACGGGAGCCTCGAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTCAGGCCCAAGTCCC 180
Ds 121 CCTCCACGGGAGCCTCGAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTCAGGCCCAAGTCCC 180
Qy 181 CGGTGGCTGTCAAGCAAGCAAGCCGGTGTATGCAAAATGGTTTCAGGCCCAAGTCCC 240
Ds 181 CGGTGGCTGTCAAGCAAGCAAGCCGGTGTATGCAAAATGGTTTCAGGCCCAAGTCCC 240
Qy 241 ATCAACAGTGGGAGCGGACCTACCTAGGCAATGGTTGGTTTGTACTTGTATGGAGAA 300
Ds 241 ATCAACAGTGGGAGCGGACCTACCTAGGCAATGGTTGGTTTGTACTTGTATGGAGAA 300
Qy 301 GCCGAGGTTTTAACTGCCGAGAGTAACTGGAAGCTGAAGAGACTTGCCTTTCGACAAAGTACA 360
Ds 301 GCCGAGGTTTTAACTGCCGAGAGTAACTGGAAGCTGAAGAGACTTGCCTTTCGACAAAGTACA 360
Qy 361 CTGGGAAACATTAACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGG 420
Ds 361 CTGGGAAACATTAACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGG 420
Qy 421 ACTGTACCTGCATCGGGGCTGGGCGAGGAGAAATGAGCTGTACCATCGCAACCGCTGCC 480
Ds 421 ACTGTACCTGCATCGGGGCTGGGCGAGGAGAAATGAGCTGTACCATCGCAACCGCTGCC 480
Qy 481 ATGAAGGGGTTCAGTCTTACCAAGATTGGTGACACCTGGAGGAGACCATGAGACTGGTG 540
Ds 481 ATGAAGGGGTTCAGTCTTACCAAGATTGGTGACACCTGGAGGAGACCATGAGACTGGTG 540
Qy 541 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAAAAGGAAATGGACCTGCAAGCCCA 600
Ds 541 GTTACATGTTAGAGTGTGTGTCTTGGTAAATGGAAAAAGGAAATGGACCTGCAAGCCCA 600
Qy 601 TAGCTGAGAAAGTGTGTATCATGCTGTGGGACTTCCCTATGTGTCTGGAGAAAGTGGG 660
Ds 601 TAGCTGAGAAAGTGTGTATCATGCTGTGGGACTTCCCTATGTGTCTGGAGAAAGTGGG 660
Qy 661 AGAAGCCTTACCAAGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Ds 661 AGAAGCCTTACCAAGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 721 GCATCACTTGCACTTCTAGAAAATAGATGCAACGATCAGGACACAAAGGACATCCTATAGAA 780
Ds 721 GCATCACTTGCACTTCTAGAAAATAGATGCAACGATCAGGACACAAAGGACATCCTATAGAA 780
Qy 781 TTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGTCTCCAGTGCATCTGCACAG 840
Ds 781 TTGGAGACACCTGGAGCAAGAGGATTAATCGAGGAAACCTGTCTCCAGTGCATCTGCACAG 840
Qy 841 GCACCGCCGAGGAGTGGAAAGTGTGAGAGGACACCTCTGTGAGAGGACATCGAGCG 900
Ds 841 GCACCGCCGAGGAGTGGAAAGTGTGAGAGGACACCTCTGTGAGAGGACATCGAGCG 900
Qy 901 GATCTGGCCCTTCAACCGATGTTCTGTGAGCTGTTTACCAACCGGAGGCTCAACCCGAGC 960
Ds 901 GATCTGGCCCTTCAACCGATGTTCTGTGAGCTGTTTACCAACCGGAGGCTCAACCCGAGC 960
Qy 961 CTCCTCCCTATGCGCACTGTGTACAGACAGTGGTGTGTCTACTCTGTGGGGATGCAGT 1020
Ds 961 CTCCTCCCTATGCGCACTGTGTACAGACAGTGGTGTGTCTACTCTGTGGGGATGCAGT 1020
Qy 1021 GGTCTGAAGACACAGGAAATAGCAAAATGCTTTTGCACGTGGCTGGGCAACGAGTGCAGT 1080
Ds 1021 GGTCTGAAGACACAGGAAATAGCAAAATGCTTTTGCACGTGGCTGGGCAACGAGTGCAGT 1080
Qy 1081 GCCAAGAGACAGCTGTGTAACCCGAGCTTACGGTGGCAACTCAAAATGGAGAGCCATGTGCT 1140
Ds 1081 GCCAAGAGACAGCTGTGTAACCCGAGCTTACGGTGGCAACTCAAAATGGAGAGCCATGTGCT 1140

Qy 1141 TACCATTACCTACAAACGACGACGACGACGACCAACTTCGAAATTAAGCAGGACCCAGA 1200
Ds 1141 TACCATTACCTACAAACGACGACGACGACGACCAACTTCGAAATTAAGCAGGACCCAGA 1200
Qy 1201 AATACTCTTTCTGCACAGACCAACATGTGTTTGGTTTTCAGACTCGAGGAGGAAATTCGAATG 1260
Ds 1201 AATACTCTTTCTGCACAGACCAACATGTGTTTGGTTTTCAGACTCGAGGAGGAAATTCGAATG 1260
Qy 1261 GTGCTTGTGCGCACTTCCCTTCTTATACAAACCAATTAACATGATTTGCACTTCTG 1320
Ds 1261 GTGCTTGTGCGCACTTCCCTTCTTATACAAACCAATTAACATGATTTGCACTTCTG 1320
Qy 1321 AGGCGAAGAGACCAACATGAAGTGGTGGGACCAACAGAACTATGATGCCAGCAGCA 1380
Ds 1321 AGGCGAAGAGACCAACATGAAGTGGTGGGACCAACAGAACTATGATGCCAGCAGCA 1380
Qy 1381 AGTTTGGGTTCTGCCCATGGTCCCAACAGGAAATCTGCAACCAATGAAGGGGTCA 1440
Ds 1381 AGTTTGGGTTCTGCCCATGGTCCCAACAGGAAATCTGCAACCAATGAAGGGGTCA 1440
Qy 1441 TGTACCCGATTGGAGATCAGTGGGATAGCAGCATGATGGGTCAATGATGAGGTGCA 1500
Ds 1441 TGTACCCGATTGGAGATCAGTGGGATAGCAGCATGATGGGTCAATGATGAGGTGCA 1500
Qy 1501 CGTGTGTGGGAATGGTCTGGGGAATGGACATGATTCCTACTCGCAGCTTCGAGATC 1560
Ds 1501 CGTGTGTGGGAATGGTCTGGGGAATGGACATGATTCCTACTCGCAGCTTCGAGATC 1560
Qy 1561 AGTCATGTTGTATGATCACTTCAATGTAAGACGACATTCACCAAGCGTCATGAAG 1620
Ds 1561 AGTCATGTTGTATGATCACTTCAATGTAAGACGACATTCACCAAGCGTCATGAAG 1620
Qy 1621 AGGGGCACTGCTGAACTGTACATGCTTCCGTGAGGCTCGGGGAGGTGGAAGTGTGATC 1680
Ds 1621 AGGGGCACTGCTGAACTGTGATCATGCTTCCGTGAGGCTCGGGGAGGTGGAAGTGTGATC 1680
Qy 1681 CCGTGCACCAATGCGAGATTCAGAGACTGCGGAGCTTTTATCAAATTTGGAGATTCATGGG 1740
Ds 1681 CCGTGCACCAATGCGAGATTCAGAGACTGCGGAGCTTTTATCAAATTTGGAGATTCATGGG 1740
Qy 1741 AGAAGTATGTCATGGTGTGAGATACCAAGTGTGCTACTGCTATGCGCGTGGCAATGGGGAGT 1800
Ds 1741 AGAAGTATGTCATGGTGTGAGATACCAAGTGTGCTACTGCTATGCGCGTGGCAATGGGGAGT 1800
Qy 1801 GGCATTCGCAACCTTTACAGACCTTATCCAGCTCAAGTGGTCTGTCGAAAGTATTTATCA 1860
Ds 1801 GGCATTCGCAACCTTTACAGACCTTATCCAGCTCAAGTGGTCTGTCGAAAGTATTTATCA 1860
Qy 1861 CTGAGACTCCGAGTCAGCCCAACTCCCAACCCATCCAGTGGAAATGACCAACAGCCATCTC 1920
Ds 1861 CTGAGACTCCGAGTCAGCCCAACTCCCAACCCATCCAGTGGAAATGACCAACAGCCATCTC 1920
Qy 1921 ACATTTCCAAAGTACATTCAGGTGGAGACTGTGAGTATCCACCCAGAAACCTTGGAT 1980
Ds 1921 ACATTTCCAAAGTACATTCAGGTGGAGACTGTGAGTATCCACCCAGAAACCTTGGAT 1980
Qy 1981 ACTGAGTCTCCTAATCTTATCAATTCGATGCTTTTTCCTGCTTTTCCAGCTTTTGAGCCAA 2040
Ds 1981 ACTGAGTCTCCTAATCTTATCAATTCGATGCTTTTTCCTGCTTTTCCAGCTTTTGAGCCAA 2040
Qy 2041 CAATCTGATTAATCTTATAGCACTTATATATTTTGTAGTGAACAAACAAATATG 2100
Ds 2041 CAATCTGATTAATCTTATAGCACTTATATATTTTGTAGTGAACAAACAAATATG 2100
Qy 2101 TGGTCAATTAATTTGACTGTGAGCTGAAAGGAAAAA 2147
Ds 2101 TGGTCAATTAATTTGACTGTGAGCTGAAAGGAAAAA 2147

RESULT 2
AX003229
LOCUS

AX003229 2147 bp DNA linear PAT 24-AUG-2000

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 03:57:22 ; Search time 199.5 Seconds

(without alignments)
11021.895 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 4096

Sequence: 1 caaacttggtgcacttgcc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlh
-Q=/cgn2.1/USPTO_spool_h/US09581651/runat_07112005_092113_27541/app_query.fasta_1.2311
-DB=UniProt -QWMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09581651 @CGN 1.1 212 @runat_07112005_092113_27541 -NCPUS=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database :

UniProt 03:.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3732.5	91.1	749	2 Q7Z391	Q7z391 homo sapien
2	3686.5	90.0	1103	2 Q6MZP4	Q6mzf4 homo sapien
3	3685.5	90.0	2357	2 Q68DT4	Q68dt4 homo sapien
4	3682.5	89.9	2444	2 Q6N025	Q6n025 homo sapien
5	3595.5	87.8	2386	1 FINC HUMAN	P02751 homo sapien
6	3592.5	87.7	2296	2 Q6N0A6	Q6n0a6 homo sapien
7	3592.5	87.7	2477	2 Q6MZU5	Q6mzu5 homo sapien
8	3588.5	87.6	2267	2 Q68DP9	Q68dp9 homo sapien
9	3567.5	87.1	2240	2 Q68DP8	Q68dp8 homo sapien
10	3393.5	82.8	2477	1 FINC MOUSE	P11276 mus musculus
11	3354.5	81.9	2477	1 FINC RAT	P04937 rattus norv
12	3328.5	81.3	2265	1 FINC BOVIN	P07589 bos taurus
13	2888	70.5	2481	2 Q6GQJ5	Q6gqa5 xenopus lae
14	2878	70.3	2481	1 FINC XENLA	Q91740 xenopus lae
15	2522	61.6	922	2 Q93405	O93405 brachydanio
16	2522	61.6	2478	2 O93406	O93406 brachydanio

17	2330	56.9	2408	2 Q6JAN2	Q6jan2 brachydanio
18	1909.5	46.6	2193	2 Q6M2M7	Q6mzm7 homo sapien
19	1496	36.5	296	2 Q8C6J7	Q8c6j7 mus musculus
20	961	23.5	190	1 FINC NOTVI	Q91400 notophthalm
21	750	18.3	141	2 Q90XQ2	Q90xq2 ambystoma m
22	543	13.3	215	2 Q6DD34	Q6dd34 xenopus lae
23	319	7.8	679	2 Q98856	Q98856 cynops pyrr
24	318	7.8	670	2 Q6DF16	Q6df16 xenopus tro
25	315.5	7.7	680	2 Q7T317	Q7t317 brachydanio
26	315	7.7	2146	2 Q68CX6	Q68cx6 homo sapien
27	313.5	7.7	673	2 Q90YB3	Q90yb3 paralicthy
28	310.5	7.6	674	2 Q98TC6	Q98tc6 cyprinus ca
29	310.5	7.6	690	2 Q9PVM5	Q9pvm5 oryzias lat
30	306	7.5	675	2 Q8QFQ6	Q8qfq6 oncorhynch
31	302.5	7.4	810	2 Q8R3F3	Q8r3f3 mus musculus
32	302	7.4	671	2 Q6PF33	Q6pf33 xenopus lae
33	302	7.4	671	2 Q9W7L6	Q9w7l6 xenopus lae
34	301	7.3	655	2 Q9W635	Q9w635 oncorhynch
35	301	7.3	661	2 Q95J44	Q95ja4 sus scrofa
36	301	7.3	662	1 MM02 RABIT	P50757 oryctolagus
37	300.5	7.3	216	2 Q99KD0	Q99kd0 mus musculus
38	300.5	7.3	686	2 Q9DE15	Q9dei5 gallus gall
39	300	7.3	662	1 MM02 MOUSE	P33434 mus musculus
40	300	7.3	662	1 MM02 RAT	P33436 rattus norv
41	300	7.3	707	1 MM09 RABIT	P41246 oryctolagus
42	300	7.3	708	1 MM09 RAT	P50282 rattus norv
43	298	7.3	412	2 Q71U44	Q71u44 bos taurus
44	298	7.3	663	1 MM02 CHICK	Q90611 gallus gall
45	297	7.3	660	1 MM02 HUMAN	P08253 homo sapien

ALIGNMENTS

RESULT 1

Q7Z391 PRELIMINARY; PRT; 749 AA.
ID Q7Z391
AC Q7Z391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp668B18150.
GN Name=DKFZp668B18150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecher H., Boscher M., Mewes H.W., Weil B., Amid C., Oanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR HSSP; Q96KP7; IFBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS01253; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Alignment Scores:
Pred. No.: 6.05e-267 Length: 749
Score: 3732.50 Matches: 656
Percent similarity: 97.19% Conservative: 0

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 06:14:58 ; Search time 42 Seconds
(without alignments)
9837.019 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 4096

Sequence: 1 caaactggtggaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Zgapop 6.0 , Zgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2 1/USPTO spool h/US09581651/runat 07112005 092114 27555/app query.fasta_1.2311
-DB=PIR -QFMT=fastan -SUPPIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09581651 @CGN 1 1 55 @runat 07112005 092114 27555 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	3579.5	87.4	2386	1 FNHU	fibronectin precu
2	3354.5	81.9	2477	2 S14428	fibronectin precu
3	3328.5	81.3	2265	1 FNBO	fibronectin - bovi
4	2888	70.5	2481	2 A43908	fibronectin - Afri
5	961	23.5	190	2 I51279	fibronectin - east
6	306	7.5	708	2 JC4364	gelatinase B (EC 3
7	301	7.3	662	2 S70355	gelatinase A (EC 3
8	300	7.3	662	2 A42496	gelatinase A (EC 3
9	300	7.3	662	2 S34780	gelatinase A (EC 3
10	300	7.3	707	1 A53796	gelatinase B (EC 3
11	300	7.3	708	2 S62907	gelatinase B (EC 3
12	298	7.3	663	1 S46492	gelatinase A (EC 3
13	297	7.3	660	1 A28153	gelatinase A (EC 3
14	294.5	7.2	712	1 I46031	gelatinase B (EC 3

ALIGNMENTS

RESULT 1

FNHU

fibronectin precursor [validated] - human

N;Alternate names: fibronectin splice form ED-A

C;Species: Homo sapiens (man)

C;Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R;Dean, D.C.; Bowlin, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656

A;Accession: A26460

A;Molecule type: DNA

A;Residues: 1-49 <DEA>

A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAAS337

R;Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A;Title: Evolution of the fibronectin gene.

A;Reference number: A26284; MUID:86111901; PMID:3003095

A;Accession: A26284

A;Molecule type: DNA

A;Residues: 1447-1540 <OLD>

A;Cross-references: GB:M12549; NID:g182688

A;Note: the authors translated the codon TTC for residue 1494 as Glu

R;Paolella, G.; Renschliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A;Reference number: S00848; MUID:88233940; PMID:3375063

A;Accession: S03917

A;Molecule type: DNA

A;Residues: 1594-1767, 'V', 1769-1783 <PAO>

A;Cross-references: EMBL:X07718; NID:g31402

A;Note: the authors translated the codon AAC for residue 1631 as Asp

R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:

A;Reference number: A24854; MUID:87030925; PMID:3770201

A;Accession: A24854

gelatinase B (EC 3
gelatinase B (EC 3
gelatinase B (EC 3
fibronectin - chic
mannose receptor p
mannose receptor p
furin (EC 3.4.21.7
zonadhesin - mouse
DELTA-like 1 - mou
phospholipase A2 r
mannose receptor,
hypothetical prote
secretory phosphol
secretory phosphol
notch protein homo
furin (EC 3.4.21.7
adhesive plaque pr
notch protein homo
crumbs protein - f
C-Delta-1 - chicke
phospholipase A2 r
notch-1 protein -
hypothetical prote
probable laminin a
transmembrane prot
MEGF8 protein - hu
phospholipase-A(2)
notch protein - fr
surface antigen - p
hypothetical prote
heparan sulfate pr

A;Molecule type: DNA
A;Residues: 1992-2147 <VTB>
A;Cross-references: GB:X04530; NID:g31436
R;Gutman, A.; Yamada, K.M.; Kornblith, A.
FEBS Lett. 207, 145-148, 1986
A;Title: Human fibronectin is synthesized as a pre-propolypeptide.
A;Reference number: A24476; MUID:87030890; PMID:3770189
A;Accession: A24476
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-14,'Q',16-38 <GUT>
R;Kornblith, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.
EMBO J. 4, 1755-1759, 1985
A;Title: Primary structure of human fibronectin: differential splicing may generate at 1
A;Reference number: A91008; MUID:85284965; PMID:2992939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080;2112-2386 <KOR>
A;Cross-references: GB:X02761
R;Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 973-2080;2112-2386 <KO2>
A;Cross-references: GB:X00739
R;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
A;Cross-references: GB:X00055; NID:g182680; PIDN:AAA52459.1; PID:g182683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386 <BER>
A;Cross-references: GB:M10905; NID:g182696; PIDN:AAA52462.1; PID:g182697
R;Umezawa, K.; Kornblith, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A;Reference number: A22245; MUID:85231203; PMID:2989004
A;Accession: A22245
A;Molecule type: mRNA
A;Residues: 1948-2067 <UMB>
A;Cross-references: GB:M27589; NID:g182705; PIDN:AAA52465.1; PID:g182706
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991;2017-2039 <UM2>
A;Cross-references: GB:M27590
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I65273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2020-2081,2113-2127 <SBK>
A;Cross-references: GB:M14060; NID:g182701; PIDN:AAA52464.1; PID:g182704
R;Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <KO3>
A;Cross-references: GB:X00799; NID:g182681; PIDN:AAA52460.1; PID:g182684

R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GARI>
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the p
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300;551-560 <GAR2>
R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GRI>
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'Q',679-703,'PT' <CAL>
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Note: residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2058 <GAR3>
R;Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080;2112-2356 <GAR4>
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
ation, and transformation.
C;Genetics:
A;Gene: GDB:FN1
A;Cross-references: GDB:119135; OMIM:135600
A;Map position: 2q34-2q34
A;Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <FHB>
F;52-87/Domain: fibronectin type I repeat homology <1F1>
F;97-135/Domain: fibronectin type I repeat homology <1F2>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 07:54:49 ; Search time 171.5 Seconds
(without alignments)
10476.095 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 4096

Sequence: 1 caacttggtggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Dgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 3735758

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092117_27656/app_query.fasta_1.2311
-DB=Published Applications_AA -QFWT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowsum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pet -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -WAP=US09581651@cgn1.1.161 @runat_07112005_092117_27656
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQUEW -NEG SCORES=0 -WAIT -DSBBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAD=10 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3664	89.5	642	16	US-10-741-601-354
2	3664	89.5	642	17	US-10-741-600-1066
3	3646.5	89.0	657	16	US-10-741-601-359
4	3646.5	89.0	657	17	US-10-741-600-1072
5	3599.5	87.9	2220	15	US-10-236-392-4
6	3599.5	87.9	2220	15	US-10-144-194A-104
7	3599.5	87.9	2355	15	US-10-447-161-3
8	3599.5	87.9	2355	16	US-10-734-564-94
9	3599.5	87.9	2355	16	US-10-491-566-104
10	3599.5	87.9	2355	17	US-10-852-335A-147
11	3599.5	87.9	2355	18	US-10-287-436A-436
12	3599.5	87.9	2355	18	US-10-287-436A-1137
13	3599.5	87.9	2355	20	US-11-040-130-28
14	3599.5	87.9	2386	16	US-10-618-281-32
15	3599.5	87.9	2474	18	US-10-450-763-52638
16	3595.5	87.8	2386	10	US-09-961-403-1
17	3595.5	87.8	2386	16	US-10-788-792-206
18	3595.5	87.8	2386	17	US-10-868-577A-59
19	3595.5	87.8	2386	18	US-10-485-758-4
20	3595.5	87.8	2386	18	US-10-485-758-9
21	3592.5	87.7	984	16	US-10-741-601-356
22	3592.5	87.7	984	17	US-10-741-600-1069
23	3592.5	87.7	2296	16	US-10-741-601-363
24	3592.5	87.7	2296	17	US-10-741-600-1075
25	3592.5	87.7	2355	16	US-10-741-601-357
26	3592.5	87.7	2355	16	US-10-741-601-366
27	3592.5	87.7	2355	17	US-10-741-600-1067
28	3592.5	87.7	2355	17	US-10-741-600-1078
29	3592.5	87.7	2386	16	US-10-741-601-360
30	3592.5	87.7	2386	17	US-10-741-600-1071
31	3448.5	84.2	2328	14	US-10-171-311-64
32	3448.5	84.2	2328	15	US-10-236-031B-70
33	3448.5	84.2	2328	15	US-10-374-979-98
34	3448.5	84.2	2328	15	US-10-182-936A-98
35	3448.5	84.2	2328	16	US-10-477-238A-677
36	3448.5	84.2	2328	16	US-10-680-287A-677
37	3448.5	84.2	2328	17	US-10-477-173-677
38	3439.5	84.0	2355	15	US-10-360-101-235
39	3424.5	83.6	2320	14	US-10-279-733-8
40	3410.5	83.3	2320	15	US-10-236-392-2
41	3296.5	80.5	1173	18	US-10-450-763-52634
42	2235	54.6	463	15	US-10-144-194A-52
43	2235	54.6	463	16	US-10-491-566-52
44	1859.5	45.4	343	9	US-09-934-706-1
45	1859.5	45.4	400	9	US-09-934-706-5

ALIGNMENTS

RESULT 1
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Alignment Scores:

Pred. No.: 2, 04e-280 Length: 642
 Score: 3664.00 Matches: 639
 Percent Similarity: 99.53% Conservative: 0
 Best Local Similarity: 99.53% Mismatches: 3
 Query Match: 89.45% Indels: 0
 DB: 16 Gaps: 0

US-09-581-651D-3 (1-2147) x US-10-741-601-354 (1-642)

QY	57	ATGCTTAGGGTCGGGGCCCGGGCTGCTGCTGCTGGCGCTCCAGTGCCTGGGACAGCG	116
DB	1	MetLeuArgGlyProGlyProGlyLeuLeuLeuLeuAlaValLeuCysLeuGlyThrAla	20
QY	117	GTGCGCTCCAGGAGCCTCGAAGAGCAAGAGGAGGCTCAGCAATGGTTCGCCCCAG	176
DB	21	ValProSerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGln	40
QY	177	TCCCGGTGGCTCAGTCAAGCAAGCCCGGTTGTTATGCAATGGGAAACACTATCAG	236
DB	41	SerProValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGln	60
QY	237	ATAAATCAACAGTGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGA	296
DB	61	IleAsnGlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGly	80
QY	297	GGAGCCGAGGTTTAACTCGGAGATAAAGCTGAAGTGAAGAGACTGCTTTCACAAAG	356
DB	81	GlySerArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLys	100
QY	357	TACACTGGGAACACTTACCGAGTGGTGCACACTATAGCGCTCAAGAGACTCCATGATC	416
DB	101	TyrThrGlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIle	120
QY	417	TGGGACTGTACTGTGATCGGGCTGGCGAGGAGAGAAATGAAGTGTACCATCGCAACCGC	476
DB	121	TrpAspCysThrCysIleGlyAlaGlyArgGlyArgGlySerCysThrIleAlaAsnArg	140
QY	477	TGCCATGAAGGGGTGAGTCTTCAAGATTGGTGCACCTGGAGAGACCATGAGACT	536
DB	141	CysHisGluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgProHisGluThr	160
QY	537	GGTGGTTACATGTTAGAGTGTGTGTTGTTGTAATGGAAGAGAGATGGACCTGCAAG	596
DB	161	GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLys	180
QY	597	CCATAGCTGAGAAGTGTGATCATGCTGCTGGGACTTCTATGTTGGTGGAGAAACG	656
DB	181	ProIleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	200
QY	657	TGGGAGAAGCCCTACCAAGGCTGGATGATGTTAGATTGTTACTGCTGGGAGAGGCAGC	716
DB	201	TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer	220
QY	717	GGAGCCATCAGTTCACCTTCTAGAATAGATCAACAGATCAGGACACAGGACATCCTAT	776
DB	221	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	240
QY	777	AGAATTGGAGACCTCGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	836
DB	241	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys	260
QY	837	ACAGCAAGCCGCGAGAGAGTGGAGTGTGAGAGGCACACCTCTGTGAGACCAATCG	896
DB	261	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	280
QY	897	ACGCGATCTGGCCCTTACCGATGTTGCTGAGCTGTTTACCAACCGCAGCCTCACCCC	956
DB	281	SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro	300
QY	957	CAGCTCTCCCTATGGCCACTGTGTACAGACAGTGGTGGTCTACTCTGTGGGGATG	1016
DB	301	GlnProProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet	320

QY	1017	CAGTGGCTGAAGACACAAAGGAAATAGCAATGCTTTGACGCTGCTGGGCAACGAGTC	1076
DB	321	GlnTrpLeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyVal	340
QY	1077	AGTGGCCAAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAATGAGAGCCATGT	1136
DB	341	SerCysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCys	360
QY	1137	GTCTTACCAATTCACCTACAGCAGCAGCAGCACTTCCGAATTTATGACGAGAC	1196
DB	361	ValLeuProPheThrTyrAsnGlyArgThrCysSerThrThrSerAsnTyrGluGlnAsp	380
QY	1197	CAGAAATACCTTCTTCACAGACACACTGTTTGGTTCAGACTCGAGAGGAGAAATCC	1256
DB	381	GlnLysTyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSer	400
QY	1257	AATGGTCTGTGTCCTTCCCTTCTATACAAACCAACCAATTCACATGATGACT	1316
DB	401	AsnGlyAlaLeuCysHisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThr	420
QY	1317	TCTGAGGCGCAGAGAGACACATGAGTGGTGGGACCAACACAGAACTATGATGCCGAC	1376
DB	421	SerGluGlyArgArgAspAsnMetLysTrpCysGlyThrGlnAsnTyrAspAlaAsp	440
QY	1377	CAGAAATTTGGTTCCTGCCCATGCTGGCCACAGGAGAAATCTGCACAAACCAATGAAAGG	1436
DB	441	GlnLysPheGlyPheCysProMetAlaAlaHisGluGluIleCysThrThrAsnGluGly	460
QY	1437	GTCAATGATCCGCAATGAGATCAGTGGGATAAGCAGCATGACATGGGTGCATGATGAGG	1496
DB	461	ValMetTyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArg	480
QY	1497	TGCACGTGTGTTGGAAATGCTGGGGAATGAGATGCTGCTGCTGCTGCTGCTGCTGCTG	1556
DB	481	CysThrCysValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArg	500
QY	1557	GATCAGTGCATGTTGTTGATGACATCATTACATGTGAACGACACATTCACCAACGCTCAT	1616
DB	501	AspGlnCysIleValAspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHis	520
QY	1617	GAAGAGGGCACATGCTGAACATGTATCATGCTTCGGTCAGGGTCGGGCGAGGTGGAAGTGT	1676
DB	521	GluGluGlyHisMetLeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCys	540
QY	1677	GATCCCTCGACCAATGCCAGATTTCAGAGACTGGGACGTTTATCAATTTGGAGATTCA	1736
DB	541	AspProValAspGlnCysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSer	560
QY	1737	TGGGAGAAGTATGTCATGTTGTCAGATACCAAGTGTCTGCTATGCGCGTGGCATTTGGG	1796
DB	561	TrpGluLysTyrValHisGlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGly	580
QY	1797	GAGTGGCATTCGCAACCTTTACAGACCTTATCAAGCTTCAAGTGTCTGCTGCTGCTGCTG	1856
DB	581	GluTrpHisCysGlnProLeuGlnThrTyrProSerSerSerGlyProValGluValPhe	600
QY	1857	ATCACTGAGACTCCGAGTCAGCCCCCACTCCCAACCCCATCCAGTGGAGTGCACCAAGCA	1916
DB	601	IleThrGluThrProSerGlnProAsnSerHisProIleGlnThrAsnAlaProGlnPro	620
QY	1917	TCTCACATTTCCAAAGTACATTTCTCAGGTGGAGACTGTGAGTATCCCAACCCCAACCTTT	1976
DB	621	SerHisIleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeu	640
QY	1977	GGATAC 1982	
DB	641	GlyTyr 642	

RESULT 2
 US-10-741-600-1066
 ; Sequence 1066, Application US/10741600
 ; Publication No. US20050026169A1
 ; GENERAL INFORMATION:

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3595.5	87.8	2386	4	US-09-061-403-1	Sequence 1, Appli
2	3579.5	87.4	2386	4	US-09-016-366A-12	Sequence 12, Appli
3	3578.5	87.4	2446	2	US-08-551-356-2	Sequence 2, Appli
4	3578.5	87.4	2446	5	PCT-US93-12687-2	Sequence 2, Appli
5	3433.5	83.8	2337	6	5455158-1	Patent No. 5455158
6	3433.5	83.8	2337	6	5455158-1	Patent No. 5455158
7	3422.5	83.6	2324	1	US-08-283-857-1	Sequence 1, Appli
8	3422.5	83.6	2324	5	PCT-US95-09819-1	Sequence 1, Appli
9	3406.5	83.2	2231	1	US-08-153-799-16	Sequence 16, Appli
10	371	9.1	65	2	US-08-982-597A-19	Sequence 19, Appli
11	371	9.1	65	3	US-09-136-218-19	Sequence 19, Appli
12	298	7.3	663	4	US-09-194-468A-30	Sequence 30, Appli

Db 1 MetLeuArgGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAla 20
Qy 117 GTGCTCCAGGAGGCTCGAAGAGGAGGAGGCTAGCAAAATGGTTTCAGCCCGCAG 176
Db 21 ValProSerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGln 40
Qy 177 TCCCGGTGGGTGTCAGTCAAGCAAGCCCGTTGTTATGACAATGGAAAACACTATCAG 236
Db 41 SerProValAlaValSerGlnSerLysProGlyCysTyrAspAenGlyLysHisTyrGln 60
Qy 237 ATAAATCAACAGTGGGAGCGACCTTACCTAGGCAATGGTTGGTTGTACTTGTATGGA 296
Db 61 IleAsnGlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGly 80
Qy 297 GGAAGCGAGGTTTAACTCGAGAGTAAACCTGAAGCTGAAGAGACTTGCTTTGACAAAG 356
Db 81 GlySerArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLys 100
Qy 357 TACACTGGAACTTACCGAGTGGTGACACTTATGAGCGTCTAAAGACTCCATGATC 416
Db 101 TyrThrGlyAenThrTyrArgValGlyAspThrTyrGluArgProLysAepSerMetIle 120
Qy 417 TGGACTCTACTGTCATCGGCGTGGGAGGAGAAATAAGCTGTACCATCGCAACCGC 476
Db 121 TrpAspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArg 140
Qy 477 TGCCATGAAGGGGTCACTCTACAGATTGGTGACACTCGAGAGACCATGAGACT 536
Db 141 CysHisGluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThr 160
Qy 537 GGTGGTACATGTTAGTGTGTCTGCTGTAATGGAAGGAGATGACCTGCCAAG 596
Db 161 GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrCysLys 180
Qy 597 CCCATAGCTGAGAAGTGTCTTTCATCATCTGCTGGGACTTCTATGTGGTGGGAGAACG 656
Db 181 ProIleAlaGlyLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 200
Qy 657 TGGGAGAACCTTACCAAGCTGGATGATGTAGATTGTCTGCTGGGAGAGGAGC 716
Db 201 TrpGlyLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 220
Qy 717 GGAGCATCACTGTCACCTTAGAATAGATGCAACGATCAGGACACAAAGGACATCCTAT 776
Db 221 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 240
Qy 777 AGAATTGGAGACACTCGGAGCAAGGATATCGAGGAAACCTCTCCAGTGCATCTGC 836
Db 241 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 260
Qy 837 ACAGGCAACGGCCGAGGAGTGGAAAGTGTGAGAGGCACTCTGTGCGAGCCACATCG 896
Db 261 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 280
Qy 897 AGCGCATCTGGCCCTTACCGCATGTCGTGCGAGCTGTTTACCAACCGCAGCTCACCCC 956
Db 281 SerGlySerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisPro 300
Qy 957 CAGCCTCTCTCTATGGCCACTGTCTCACAGACAGTGGTGGTCTACTCTGTGGGGATG 1016
Db 301 GlnProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet 320
Qy 1017 CAGTGGCTGAGACACAGAGAAATAGCAATGCTTTGCACTGCTGGGCAACGGAGTC 1076
Db 321 GlnTrpLeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyVal 340
Qy 1077 AGCTCCCAAGACAGACTGTAAACCACTTACCGTGGCAACTCAATGGAGAGCCATGT 1136
Db 341 SerCysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCys 360
Qy 1137 GTCTTACCATTACCTACCAACGACAGGAGC----- 1166
Db 361 ValLeuProPheThrTyrAsnGlyArgThrPheTyrSerCysThrThrGluGlyArgGln 380

Qy 1167 -----GACAGCACAACTTCGAATTATGAGCAGGACGACAGAAATACTCTTTC 1211
Db 381 AspGlyHisLeuTrpCysSerThrThrSerAsnTyrGlnGlnAspGlnLysTyrSerPhe 400
Qy 1212 TGACAGACACACTGTTTTGGTTTCAGACTCGAGGAGAAATTCGAATGTGCTTGTGC 1271
Db 401 CysThrAspHisThrValLeuValGlnThrGlnGlyGlyAsnSerAsnGlyAlaLeuCys 420
Qy 1272 CACTTCCCTTCTATACAAACCAACCAATTACACTGATTGCCTTCTGAGGCGCAGAGA 1331
Db 421 HisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArg 440
Qy 1332 GACAACTAAGTGGTGGGACACACAGAACTATGATGCCACCAAGAGTTTGGGTTTC 1391
Db 441 AspAsnMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPhe 460
Qy 1392 TGCCCCATGGCTGCCACGAGGAAATCTGCACAAACCAATCAAGGGGTTCATGTACCGCAT 1451
Db 461 CysProMetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIle 480
Qy 1452 GGAGATCAGTGGGATTAAGCAGCATGATCGGTTCACATGATGAGGTGACGTGTGTGGG 1511
Db 481 GlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGly 500
Qy 1512 AATGGTCTGGGAAATGGACATGCATTGCCCTTACTCGCAGCTTCGAGATCAGTGCATTGTT 1571
Db 501 AsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleVal 520
Qy 1572 GATGACATCACTTACAAATGTGAACGACACATTCACAAAGCGTCATGAAGAGGGGCACATG 1631
Db 521 AspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMet 540
Qy 1632 CTGAACCTGATCCTTCGGTTCAGGCTCGGGCGGAGGTGGAACTGTGATCCGTCGACCAA 1691
Db 541 LeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGln 560
Qy 1692 TGCCAGGATTCAGAGACTGGGAGCTTTTATCAATTCGAGATTCATGGGAGAGTATGTG 1751
Db 561 CysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrVal 580
Qy 1752 CATGGTTCAGATACCACTGCTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCAA 1811
Db 581 HisGlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGln 600
Qy 1812 CTTTACAGACTATCCAAGCTCAAGTGGTCTCTGTCGAAGTATTTATCCTGAGACTCCG 1871
Db 601 ProLeuGlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrPro 620
Qy 1872 AGTCAGCCCACTCCCAACCCCATCCAGTGAATGCACACAGCCATCTCACATTTTCCAAG 1931
Db 621 SerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLys 640
Qy 1932 TACATTCCTCAGGTGGAGACCT 1952
Db 641 TyrIleLeuArgTrpArgPro 647
RESULT 2
US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 11, 2005, 03:51:06 ; Search time 163.5 seconds
(without alignments)
10157.487 Million cell updates/sec

Title: US-09-581-651D-3

Perfect score: 4096

Sequence: 1 caacttggtggcaacttgc.....aaaaaaaaaaaaaaaaaaaaa 2147

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_b/US09581651/runat_07112005_092112_27535/app_query.fasta_1.2311
-DB=A_Geneseq -QFWT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=dbs -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CGN 1.1 163 @runat_07112005_092112_27535 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3781	92.3	660	2 AAY28901	Aay28901 Human mig
2	3693.5	90.2	2182	8 ADR66462	Adr66462 Human pro
3	3693.5	90.2	2182	8 ADR66120	Adr66120 Human pro
4	3680.5	89.9	720	2 AAY28914	Aay28914 Fibronect
5	3680	89.8	642	8 ADR67316	Adr67316 Human bla
6	3680	89.8	642	8 ADS17489	Adsl17489 Amino aci
7	3680	89.8	642	8 ADR97658	Adr97658 Human fib
8	3664	89.5	642	8 ADQ39403	Adq39403 Human myo
9	3658	89.3	2447	4 AAM40434	Aam40434 Human pol
10	3658	89.3	2447	4 AAM40435	Aam40435 Human pol

11	3658	89.3	2447	4	AAM40433	Aam40433 Human pol
12	3658	89.3	2447	4	AAM40432	Aam40432 Human pol
13	3646.5	89.0	657	8	ADQ39409	Adq39409 Human myo
14	3599.5	87.9	2220	6	ABO01289	Ab001289 Human pro
15	3599.5	87.9	2266	6	ABR40124	AbR40124 Human cel
16	3599.5	87.9	2355	6	ABR58335	AbR58335 NM_00202
17	3599.5	87.9	2355	7	ADP65196	Adp65196 Human fib
18	3599.5	87.9	2355	8	ADG89560	Adg89560 Human fib
19	3599.5	87.9	2355	8	ADO55175	Ado55175 Protein #
20	3599.5	87.9	2355	8	ADQ26085	Adq26085 Fibronect
21	3599.5	87.9	2355	8	ADQ29668	Adq29668 Human col
22	3599.5	87.9	2355	8	ADR67315	Adr67315 Human bla
23	3599.5	87.9	2386	8	ADO55174	Ado55174 Protein #
24	3599.5	87.9	2474	4	ABG22279	Abg22279 Novel hum
25	3595.5	87.8	2386	5	AAO17353	Aao17353 Human fib
26	3595.5	87.8	2386	6	ABR81866	AbR81866 Human fib
27	3595.5	87.8	2386	7	ADD18770	Add18770 Human dia
28	3595.5	87.8	2386	7	ADG63324	Adg63324 Human pro
29	3595.5	87.8	2386	7	ADR92200	Adr92200 Fibronect
30	3593.5	87.7	2446	3	AAB50377	Aab50377 Human fib
31	3592.5	87.7	984	8	ADQ39406	Adq39406 Human myo
32	3592.5	87.7	2265	4	AAM38647	Aam38647 Human pol
33	3592.5	87.7	2296	8	ADQ39412	Adq39412 Human myo
34	3592.5	87.7	2330	4	AAM38646	Aam38646 Human pol
35	3592.5	87.7	2355	4	AAM38649	Aam38649 Human pol
36	3592.5	87.7	2355	8	ADQ39415	Adq39415 Human myo
37	3592.5	87.7	2355	8	ADQ39404	Adq39404 Human myo
38	3592.5	87.7	2386	4	AAM38648	Aam38648 Human pol
39	3592.5	87.7	2386	8	ADQ39408	Adq39408 Human myo
40	3586.5	87.6	2477	2	AAM95955	Aam95955 Human fib
41	3584.5	87.5	2220	8	ADN95950	Adn95950 Human NOV
42	3579.5	87.4	2386	2	AAM63171	Aam63171 Amino aci
43	3578.5	87.4	2446	2	AAR60021	Aar60021 Fibronoge
44	3546	86.6	2352	8	ADK00410	Adk00410 Fibronect
45	3448.5	84.2	2328	4	AAG68182	Aag68182 Fibronect

ALIGNMENTS

RESULT 1

AAY28901
ID AAY28901 standard; protein; 660 AA.

XX AC AAY28901;

XX XX 21-SEP-1999 (first entry)

XX DE Human migration stimulating factor (MSF) 1-alpha protein.

XX KW Migration stimulatory factor; MSF; cell migration; modulation; human; wound healing; scarring; MSF1-alpha.

XX OS Homo sapiens.

XX XX WO9931233-A1.

XX XX 24-JUN-1999.

XX PF 15-DEC-1998; 98WO-GB003766.

XX PR 16-DEC-1997; 97GB-00026539.

XX XX (UYDU-) UNIV DUNDEE.

XX XX Schor SL, Schor AM;

XX XX WPI; 1999-430039/36.

XX DR N-PSDE; AAX81299.

XX XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX PS Claim 1; Page 53; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. The present sequence
CC represents the human MSF1-alpha protein

XX
SQ Sequence 660 AA;

Alignment Scores:

Pred. No.:	2.08e-293	Length:	660
Score:	3781.00	Matches:	660
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	92.31%	Indels:	0
DB:	2	Gaps:	0

US-09-581-651D-3 (1-2147) x AAY28901 (1-660)

Qy	3	AAC	TG	TG	TG	GCA	ACT	TG	CCT	CCC	GT	GCG	GG	CG	TCT	CCCC	CC	CG	TCT	CA	CA	TG	CTT	62
Db	1	Asn	Leu	Val	Ala	Thr	Cys	Leu	Pro	Val	Arg	Ala	Ser	Leu	Pro	His	Arg	Leu	Asn	Met	Leu	20		
Qy	63	AGG	GT	CCG	GG	CCG	CG	TG	CT	GT	CT	G	TG	CC	GT	CC	AG	TG	CG	T	GG	GC	122	
Db	21	Arg	Gly	Pro	Gly	Pro	Gly	Leu	Leu	Leu	Ala	Val	Gln	Cys	Leu	Gly	Thr	Ala	Val	Pro	40			
Qy	123	TCC	AC	GGG	AG	CG	TCC	GA	AG	CA	AG	CG	CT	CAG	CA	AA	TG	CT	CAG	CCC	CG	182		
Db	41	Ser	Thr	Gly	Ala	Ser	Leu	Ser	Leu	Ser	Leu	Arg	Gln	Ala	Gln	Met	Val	Gln	Pro	Gln	Ser	60		
Qy	183	GTG	CT	GT	CAG	TCA	AG	CA	AG	CCG	CG	TG	TG	TAT	TGA	CA	AT	TG	GA	AA	CA	242		
Db	61	Val	Ala	Val	Ser	Gln	Ser	Leu	Ser	Pro	Gly	Cys	Tyr	Asp	Asn	Gly	Leu	His	Tyr	Gln	Leu	80		
Qy	243	CAA	CA	GT	GGG	AG	CG	GA	CT	CA	CT	AG	CA	AT	TG	CT	GG	TG	TG	TG	TG	302		
Db	81	Gln	Gln	Trp	Glu	Arg	Thr	Thr	Leu	Gly	Asn	Ala	Leu	Val	Cys	Thr	Cys	Tyr	Gly	Gly	Ser	100		
Qy	303	CGA	GT	TTT	TAA	CTG	CG	AG	AG	TAA	AA	CTG	AA	GC	TG	AG	AG	AC	TG	TTT	CG	362		
Db	101	Arg	Gly	Phe	Asn	Cys	Glu	Ser	Leu	Pro	Glu	Ala	Glu	Glu	Thr	Cys	Phe	Asp	Leu	Ser	Thr	120		
Qy	363	GGG	AA	CA	CT	TAC	CG	AG	TGG	GT	GAC	ACT	TAT	GAG	CG	TCT	TAA	AG	ACT	CC	CA	422		
Db	121	Gly	Asn	Thr	Tyr	Arg	Val	Gly	Asp	Thr	Tyr	Glu	Arg	Pro	Leu	Asp	Ser	Met	Leu	Trp	Asp	140		
Qy	423	TGT	AC	TG	CTG	AT	CGG	GT	GGG	CG	AGG	AG	AAT	AA	GCT	GT	AC	CA	TG	CC	AA	482		
Db	141	Cys	Thr	Cys	Ile	Gly	Ala	Gly	Arg	Gly	Arg	Gly	Ser	Cys	Thr	Leu	Ala	Asn	Arg	Cys	His	160		
Qy	483	GN	AG	GGG	GT	CA	CT	TAC	AA	GA	TGG	TG	AC	CTG	GAG	AG	CA	CC	AT	CAG	AG	542		
Db	161	Glu	Gly	Gly	Gln	Ser	Tyr	Ile	Gly	Asp	Thr	Trp	Arg	Pro	His	Gln	Thr	Gly	Gly	180				
Qy	543	TAC	AT	GTT	TAG	AG	TG	TG	TG	TCT	TG	TAT	TG	AAA	AG	AGA	TGG	AC	CT	CG	AA	602		
Db	181	Tyr	Met	Leu	Glu	Cys	Val	Cys	Leu	Gly	Asn	Gly	Leu	Gly	Glu	Trp	Thr	Cys	Leu	Pro	200			
Qy	603	GCT	G	AA	AG	ACT	GTT	TG	AT	CA	TG	CT	GTG	GC	ACT	TCT	TAT	TG	TG	TCG	GA	662		
Db	201	Ala	Glu	Leu	Cys	Phe	Asp	His	Ala	Ala	Gly	Thr	Ser	Tyr	Val	Val	Gly	Glu	Thr	Trp	Glu	220		
Qy	663	AG	CC	CT	AC	CA	GG	CT	GG	AT	GAT	TG	GT	TAG	AT	TG	CT	GG	AG	AG	CG	722		
Db	221	Lys	Pro	Tyr	Gln	Gly	Trp	Met	Met	Val	Asp	Cys	Thr	Cys	Leu	Gly	Glu	Ser	Gly	Arg	240			
Qy	723	AT	C	ACT	TTG	CAC	TCT	T	AGA	AAT	TAG	AT	G	CA	C	GA	T	CAG	CA	CA	AG	7		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:57:14 ; Search time 202.745 Seconds
(without alignments)
1666.981 Million cell updates/sec

Title: US-09-581-651D-1
Perfect score: 3781
Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILRWRPVSIPRNLGY 660

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3732.5	98.7	749	2 Q7Z391	Q7Z391 homo sapien
2	3686.5	97.5	1103	2 Q6MZFA	Q6mzf4 homo sapien
3	3685.5	97.5	2357	2 Q68DT4	Q68dt4 homo sapien
4	3682.5	97.4	2444	2 Q6N025	Q6n025 homo sapien
5	3595.5	95.1	2386	1 F1NC_HUMAN	P02751 homo sapien
6	3592.5	95.0	2296	2 Q6N0A6	Q6n0a6 homo sapien
7	3592.5	95.0	2477	2 Q6MZU5	Q6mzu5 homo sapien
8	3588.5	94.9	2267	2 Q68DP9	Q68dp9 homo sapien
9	3567.5	94.4	2240	2 Q68DP8	Q68dp8 homo sapien
10	3393.5	89.8	2477	1 F1NC_MOUSE	P11276 mus musculus
11	3354.5	88.7	2477	1 F1NC_RAT	P04937 rattus norv
12	3328.5	88.0	2265	1 F1NC_BOVIN	P07589 bos taurus
13	2888	76.4	2481	2 Q6GQAS	Q6gqa5 xenopus lae
14	2878	76.1	2481	1 F1NC_XENLA	Q91740 xenopus lae
15	2513.5	66.5	922	2 Q93405	O93405 brachydanio
16	2513.5	66.5	2478	2 O93406	O93406 brachydanio
17	2328	61.6	2408	2 Q6JAN2	Q6jan2 brachydanio
18	1909.5	50.5	2193	2 Q6CMZ7	Q6cmz7 homo sapien
19	1496	39.6	296	2 Q8C6J7	Q8c6j7 mus musculus
20	961	25.4	190	1 F1NC_NOTVI	Q91400 notophthala
21	750	19.8	141	2 Q90XQ2	Q90xq2 ambystoma m
22	541	14.3	215	2 Q6DD34	Q6dd34 xenopus lae
23	317.5	8.4	679	2 Q98856	Q98856 cynops pyrr
24	313.5	8.3	673	2 Q90YB3	Q90yb3 paralicthy
25	310.5	8.2	690	2 Q9PVM5	Q9pvm5 oryzias lat
26	307	8.1	670	2 Q6DFL6	Q6df16 xenopus tro
27	307	8.1	680	2 Q7T317	Q7t317 brachydanio
28	306	8.1	675	2 Q8QF06	Q8qf06 oncorhynch
29	302.5	8.0	810	2 Q8R3F3	Q8r3f3 mus musculus
30	302	8.0	671	2 Q6FF33	Q6pf33 xenopus lae
31	302	8.0	671	2 Q9W7L6	Q9w7l6 xenopus lae

32 301 8.0 655 2 Q9W635 oncorhynchu
33 301 8.0 662 1 MM02_RABIT P50757 oryctolagus
34 300.5 7.9 216 2 Q99KD0 Q99kd0 mus musculu
35 300.5 7.9 674 2 Q98TC6 Q98tc6 cyprinus ca
36 300 7.9 662 1 MM02_MOUSE P33434 mus musculu
37 300 7.9 662 1 MM02_RAT P33436 rattus norv
38 300 7.9 707 1 MM09_RABIT P41246 oryctolagus
39 299.5 7.9 2146 2 Q68CX6 Q68cx6 homo sapien
40 298 7.9 412 2 Q71U44 Q71u44 bos taurus
41 298 7.9 663 1 MM02_CHICK Q90611 gallus gall
42 297 7.9 660 1 MM02_HUMAN P08253 homo sapien
43 296.5 7.8 1256 1 F1NC_CHICK P11722 gallus gall
44 296 7.8 632 2 Q9N1P6 Q9n1p6 canis famil
45 295 7.8 211 2 Q7L553 Q7l553 homo sapien

ALIGNMENTS

RESULT 1
Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecher H., Boecher M., Mewes H.W., Weill B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -
DR HSP; Q96KP7; 1PBR.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 98.7%; Score 3732.5; DB 2; Length 749;

Best Local Similarity 97.2%; Pred. No. 2.6e-261;

Matches 656; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy 1 NLVATCLPVRASLPHRLNMLRGPGGLLLAVQCLGTAVPSTGASKSKRQAOQWVQPS 60
Db 75 NLVATCLPVRASLPHRLNMLRGPGGLLLAVLCLGTAVPSTGASKSKRQAOQWVQPS 134
Qy 61 VAVSQSGPGCVNKGKHYQINQWERTYLGVALVCTCYGGSGRGNFNCSEKPEAEETCFDKYT 120
Db 135 VAVSQSGPGCVNKGKHYQINQWERTYLGVALVCTCYGGSGRGNFNCSEKPEAEETCFDKYT 194
Qy 121 GNTYRVGDTYERPKDSMIWDCITCIGAGRGRIISCTIANRCHGGQSYKIGDTWRPHETGG 180
Db 195 GNTYRVGDTYERPKDSMIWDCITCIGAGRGRIISCTIANRCHGGQSYKIGDTWRPHETGG 254
Qy 181 YMLECVLGNKGKGBWTCKPTAEKCFDHAAGTSYVVGITWEKPYQGMWVDTCTCLGEGSGR 240

Db 255 YMLECVCLGNGKEWTKPIAEKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGR 314
Qy 241 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 300
Db 315 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 374
Qy 301 SGPTFDVRAAYVQPPHPPYGHCVTDGVTYVSGWQWLKTOGNKQMLCTCLGNGVSC 360
Db 375 SGPTFDVRAAYVQPPHPPYGHCVTDGVTYVSGWQWLKTOGNKQMLCTCLGNGVSC 434
Qy 361 QETAVTQTYGNSNGEPCVLPFTYNDRT-----DSTTSNYEQDQKYSFCT 405
Db 435 QETAVTQTYGNSNEEPCVLPFTYNGRTFYSCCTTEGRODGLWCSTTSNYEQDQKYSFCT 494
Qy 406 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 465
Db 495 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 554
Qy 466 MAHEEICTTNEGVYRIGDQWDKQHDGMHMRCTCVNGRGEWTCIAYSQLRDQCI VDD 525
Db 555 MAHEEICTTNEGVYRIGDQWDKQHDGMHMRCTCVNGRGEWTCIAYSQLRDQCI VDD 614
Qy 526 ITYNVNDTFHKKRHEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIIGDSWEKYVHG 585
Db 615 ITYNVNDTFHKKRHEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIIGDSWEKYVHG 674
Qy 586 VRYQCYCYGRGIGEWHCOPLOTYPSSSGPVEVFITETPSQNSHPIONNAPQSHISKYI 645
Db 675 VRYQCYCYGRGIGEWHCOPLOTYPSSSGPVEVFITETPSQNSHPIONNAPQSHISKYI 734
Qy 646 LRWRPVSIPPRNLGY 660
Db 735 LRWRPVSIPPRNLGY 749

RESULT 2
Q6MZP4
ID Q6MZP4 PRELIMINARY; PRT; 1103 AA.
AC Q6MZP4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219)
DE (Fragment)
GN Name=DKFZp686K139; Synonyms=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG The German Human cDNA Consortium;
RA Ansgore W., Krieger S., Regier T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amlid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BX649182; CAB46200.1; -;
DR EMBL; BX640802; CAB45885.1; -;
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR00562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR000112; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.

DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82PEC4CAF634AD56 CRC64;
Query Match 97.5%; Score 3686.5; DB 2; Length 1103;
Best Local Similarity 97.3%; Pred. No. 8.2e-258;
Matches 647; Conservative 0; Mismatches 3; Indels 15; Gaps 1;
Qy 1 NLVATCLPVRASLPHRLNMLRGPGGLLLAVQCLGTAVPSTGASKRQAAQVQPSQ 60
Db 72 NLVATCLPVRASLPHRLNMLRGPGGLLLAVLCLGTAVPSTGASKRQAAQVQPSQ 131
Qy 61 VAVSQSPGCVYDNGKHQYQINQWERTYLGNAVCTCYGSGRGFNCESKPEABETCFDKYT 120
Db 132 VAVSQSPGCVYDNGKHQYQINQWERTYLGNAVCTCYGSGRGFNCESKPEABETCFDKYT 191
Qy 121 GNTYRVGDTYERPKDSIMWDCTCIGAGRGISCTIANRCHGGOSYKIGDTWRPHETGG 180
Db 192 GNTYRVGDTYERPKDSIMWDCTCIGAGRGISCTIANRCHGGOSYKIGDTWRPHETGG 251
Qy 181 YMLECVCLGNGKEWTKPIAEKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGR 240
Db 252 YMLECVCLGNGKEWTKPIAEKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGR 311
Qy 241 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 300
Db 312 ITCTSRNRCDQDTRTSYRIGDTSKDNKGNLLQICITGNGRGEWKCERHTSVQTTSSG 371
Qy 301 SGPTFDVRAAYVQPPHPPYGHCVTDGVTYVSGWQWLKTOGNKQMLCTCLGNGVSC 360
Db 372 SGPTFDVRAAYVQPPHPPYGHCVTDGVTYVSGWQWLKTOGNKQMLCTCLGNGVSC 431
Qy 361 QETAVTQTYGNSNGEPCVLPFTYNDRT-----DSTTSNYEQDQKYSFCT 405
Db 432 QETAVTQTYGNSNGEPCVLPFTYNGRTFYSCCTTEGRODGLWCSTTSNYEQDQKYSFCT 491
Qy 406 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 465
Db 492 DHTVLVQTRGNSNGALCHFFPFLYNNHNYTCTSEGRDNNKWCCTTQNYDADQKFGFCP 551
Qy 466 MAHEEICTTNEGVYRIGDQWDKQHDGMHMRCTCVNGRGEWTCIAYSQLRDQCI VDD 525
Db 552 MAHEEICTTNEGVYRIGDQWDKQHDGMHMRCTCVNGRGEWTCIAYSQLRDQCI VDD 611
Qy 526 ITYNVNDTFHKKRHEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIIGDSWEKYVHG 585
Db 612 ITYNVNDTFHKKRHEEGHMLNCTCFQGRGRWKCDPVDQCDSETGTFTFYIIGDSWEKYVHG 671
Qy 586 VRYQCYCYGRGIGEWHCOPLOTYPSSSGPVEVFITETPSQNSHPIONNAPQSHISKYI 645
Db 672 VRYQCYCYGRGIGEWHCOPLOTYPSSSGPVEVFITETPSQNSHPIONNAPQSHISKYI 731
Qy 646 LRWRP 650
Db 732 LRWRP 736

RESULT 3
Q68DT4
ID Q68DT4 PRELIMINARY; PRT; 2357 AA.
AC Q68DT4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp686F10164.
GN Name=DKFZp686F10164;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:12:29 ; Search time 50.0392 Seconds
(without alignments)
1269.066 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781

Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILRWKPVSPRNLGY 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3579.5	94.7	2386	1 FNHU	fibronectin precu
2	3354.5	88.7	2477	2 S14428	fibronectin precu
3	3328.5	88.0	2265	1 FNBO	fibronectin - bovi
4	2888	76.4	2481	2 A43908	fibronectin - Afri
5	961	25.4	190	2 I51279	fibronectin - east
6	302.5	8.0	708	2 JC4364	gelatinase B (EC 3
7	301	8.0	662	2 S70365	gelatinase A (EC 3
8	300	7.9	662	2 A42496	gelatinase A (EC 3
9	300	7.9	662	2 S34780	gelatinase A (EC 3
10	300	7.9	707	1 A53796	gelatinase B (EC 3
11	298	7.9	663	1 S46492	gelatinase A (EC 3
12	297	7.9	660	1 A28153	gelatinase A (EC 3
13	294.5	7.8	708	2 S62907	gelatinase B (EC 3
14	294.5	7.8	712	1 I46031	gelatinase B (EC 3
15	294.5	7.8	730	1 I52580	gelatinase B (EC 3
16	294.5	7.8	730	2 JC1456	gelatinase B (EC 3
17	285.5	7.6	707	1 A34458	gelatinase B (EC 3
18	228.5	6.0	1020	2 A29355	fibronectin - chic
19	208	5.5	1455	1 A48925	mannose receptor p
20	199.5	5.3	1456	1 A36563	mannose receptor p
21	199	5.3	5376	2 T42215	mannose receptor,
22	190.5	5.0	1479	2 T42210	mannose receptor,
23	188.5	5.0	1584	2 T22674	hypothetical prote
24	188	5.0	1458	1 A49707	phospholipase A2 r
25	185.5	4.9	1326	2 B56395	secretory phosphol
26	185.5	4.9	1465	2 A56395	secretory phosphol
27	182	4.8	722	2 I48324	DELTA-like 1 - mou
28	181.5	4.8	473	2 A56175	adhesive plaque pr
29	179.5	4.7	1463	2 A53210	phospholipase A2 r

ALIGNMENTS

RESULT 1

FNHU

fibronectin precursor [validated] - human
N;Alternate names: fibronectin splice form ED-A

C;Species: Homo sapiens (man)

C;Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R;Dean, D.C.; Bowlus, C.L.; Bourgeois, S

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656

A;Accession: A26460

A;Molecule type: DNA

A;Residues: 1-49 <DE>

A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:G182686; PIDN:AA5337

R;Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A;Title: Evolution of the fibronectin gene.

A;Reference number: A26284; MUID:86111901; PMID:3003095

A;Accession: A26284

A;Molecule type: DNA

A;Residues: 1447-1540 <OLD>

A;Cross-references: GB:M12549; NID:G182688

R;Paolella, G.; Henschcliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A;Reference number: S00848; MUID:88233940; PMID:3375063

A;Accession: S03917

A;Molecule type: DNA

A;Residues: 1594-1767, 'V', 1769-1783 <PAO>

A;Cross-references: EMBL:X07719; NID:G31402

A;Note: the authors translated the codon AAC for residue 1631 as Asp

R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:

A;Reference number: A24854; MUID:87030929; PMID:3770201

A;Accession: A24854

A;Molecule type: DNA

A;Residues: 1992-2147 <VIB>

A;Cross-references: GB:X04530; NID:G31436

R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.

FEBS Lett. 207, 145-148, 1986

A;Title: Human fibronectin is synthesized as a pre-propolypeptide.

A;Reference number: A24476; MUID:87030890; PMID:3770189

A;Accession: A24476

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-14, 'Q', 16-38 <GUT>

R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.

EMBO J. 4, 1755-1759, 1985

A;Title: Primary structure of human fibronectin: differential splicing may generate at 1

notch protein - fr
thrombospondin 1 p
C-Delta-1 - chicke
Xotch protein - Af
notch-1 protein -
phospholipase-A(2)
notch protein homo
crumbs protein - f
hypothetical prote
probable laminin a
coagulation factor
laminin alpha-1 ch
insulin-like growt
cation-independent
cation-independent
transmembrane prot

A;Reference number: A91008; MUID:85284965; PMID:2992939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080;2112-2386 <KOR>
A;Cross-references: GB:X02761
R;Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptide
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 973-2080;2112-2386 <KO2>
A;Cross-references: GB:X00739
R;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
A;Cross-references: GB:K00055; NID:G182680; PIDN:AAA52459.1; PID:G182683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386

A;Cross-references: GB:M10905; NID:G182696; PIDN:AAA52462.1; PID:G182697
R;Umezawa, K.; Kornblith, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A;Reference number: A22245; MUID:85231203; PMID:2989004
A;Accession: A22245
A;Molecule type: mRNA
A;Residues: 1948-2067 <UWE>
A;Cross-references: GB:M27589; NID:G182705; PIDN:AAA52465.1; PID:G182706
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991;2017-2039 <UM2>
A;Cross-references: GB:M27590
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different me
A;Reference number: I52394; MUID:87026578; PMID:3021206
A;Accession: I65273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2020-2081,2113-2127 <SEK>
A;Cross-references: GB:M4060; NID:G182701; PIDN:AAA52464.1; PID:G182704
R;Kornblith, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectins
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <K03>
A;Cross-references: GB:K00799; NID:G182681; PIDN:AAA52460.1; PID:G182684
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GAR1>
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the F
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300;551-560 <GAR2>

R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GRI>
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Sirl, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human p1
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'Q',679-703,'PT' <CAL>
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Note: residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2058 <GAR3>
R;Tresselt, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080;2112-2356 <GAR4>
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
action, and transformation.
C;Genetics:
A;Gene: GDB:FNI
A;Cross-references: GDB:I19135; OMIM:135600
A;Map position: 2q34-2q34
A;Inrons: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibronectin type I repeat homology <FHB>
F;52-87/Domain: fibronectin type I repeat homology <1F1>
F;97-135/Domain: fibronectin type I repeat homology <1F2>
F;141-179/Domain: fibronectin type I repeat homology <1F3>
F;186-225/Domain: fibronectin type I repeat homology <1F4>
F;231-270/Domain: fibronectin type I repeat homology <1F5>
F;308-608/Domain: collagen binding <CBR>
F;308-342/Domain: fibronectin type I repeat homology <1F6>
F;360-401/Domain: fibronectin type II repeat homology <2F1>
F;420-461/Domain: fibronectin type II repeat homology <2F2>
F;470-508/Domain: fibronectin type I repeat homology <1F7>
F;518-555/Domain: fibronectin type I repeat homology <1F8>
F;561-599/Domain: fibronectin type I repeat homology <1F9>
F;603-692/Domain: fibronectin type III repeat homology <3FA>
F;616-706/Domain: heparin binding <HPB>
F;719-801/Domain: fibronectin type III repeat homology <3FB>
F;810-891/Domain: fibronectin type III repeat homology <3FC>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 08:29:55 ; Search time 223.451 Seconds
(without alignments)
1235.843 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781

Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILWRPVSIPPRNLGY 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB. pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW PUB. pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB. pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB. pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB. pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB. pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB. pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB. pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB. pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB. pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB. pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB. pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB. pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB. pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB. pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB. pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB. pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB. pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB. pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW PUB. pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB. pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB. pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3664	96.9	642	16 US-10-741-601-354	Sequence 354, App
2	3664	96.9	642	17 US-10-741-600-1066	Sequence 1066, App
3	3646.5	96.4	657	16 US-10-741-601-359	Sequence 359, App
4	3646.5	96.4	657	17 US-10-741-600-1072	Sequence 1072, App
5	3599.5	95.2	2220	15 US-10-236-392-4	Sequence 4, Appl1
6	3599.5	95.2	2355	15 US-10-144-194A-104	Sequence 104, App
7	3599.5	95.2	2355	15 US-10-447-161-3	Sequence 3, Appl1
8	3599.5	95.2	2355	16 US-10-734-564-94	Sequence 94, Appl1
9	3599.5	95.2	2355	16 US-10-491-566-104	Sequence 104, App
10	3599.5	95.2	2355	17 US-10-852-335A-147	Sequence 147, App
11	3599.5	95.2	2355	18 US-10-287-436A-436	Sequence 436, App

12	3599.5	95.2	2355	18	US-10-287-436A-1137	Sequence 1137, App
13	3599.5	95.2	2355	20	US-11-040-130-28	Sequence 28, Appl1
14	3599.5	95.2	2386	16	US-10-618-281-32	Sequence 32, Appl1
15	3599.5	95.2	2474	18	US-10-450-763-52638	Sequence 52638, A
16	3595.5	95.1	2386	10	US-09-961-403-1	Sequence 1, Appl1
17	3595.5	95.1	2386	16	US-10-788-792-206	Sequence 206, App
18	3595.5	95.1	2386	17	US-10-868-577A-59	Sequence 59, Appl1
19	3595.5	95.1	2386	18	US-10-485-758-4	Sequence 4, Appl1
20	3595.5	95.1	2386	18	US-10-485-758-9	Sequence 9, Appl1
21	3592.5	95.0	984	16	US-10-741-601-356	Sequence 356, App
22	3592.5	95.0	984	17	US-10-741-600-1069	Sequence 1069, App
23	3592.5	95.0	2296	16	US-10-741-601-363	Sequence 363, App
24	3592.5	95.0	2296	17	US-10-741-600-1075	Sequence 1075, App
25	3592.5	95.0	2355	16	US-10-741-601-357	Sequence 357, App
26	3592.5	95.0	2355	16	US-10-741-601-366	Sequence 366, App
27	3592.5	95.0	2355	17	US-10-741-600-1067	Sequence 1067, App
28	3592.5	95.0	2355	17	US-10-741-600-1078	Sequence 1078, App
29	3592.5	95.0	2386	16	US-10-741-601-360	Sequence 360, App
30	3592.5	95.0	2386	16	US-10-741-600-1071	Sequence 1071, App
31	3448.5	91.2	2328	14	US-10-171-311-64	Sequence 64, Appl1
32	3448.5	91.2	2328	15	US-10-236-031B-70	Sequence 70, Appl1
33	3448.5	91.2	2328	15	US-10-374-979-98	Sequence 98, Appl1
34	3448.5	91.2	2328	15	US-10-182-936A-98	Sequence 98, Appl1
35	3448.5	91.2	2328	16	US-10-477-238A-677	Sequence 677, App
36	3448.5	91.2	2328	16	US-10-680-287A-677	Sequence 677, App
37	3448.5	91.2	2328	17	US-10-477-173-677	Sequence 677, App
38	3439.5	91.0	2355	15	US-10-360-101-235	Sequence 235, App
39	3424.5	90.6	2320	14	US-10-379-733-8	Sequence 8, Appl1
40	3410.5	90.2	2320	15	US-10-236-392-2	Sequence 2, Appl1
41	3286.5	87.2	1173	18	US-10-450-763-52634	Sequence 52634, A
42	2144	56.7	463	15	US-10-144-194A-52	Sequence 52, Appl1
43	2144	56.7	463	16	US-10-491-566-52	Sequence 52, Appl1
44	1859.5	49.2	343	9	US-09-934-706-1	Sequence 1, Appl1
45	1859.5	49.2	400	9	US-09-934-706-5	Sequence 5, Appl1

ALIGNMENTS

RESULT 1

US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 96.9%; Score 3664; DB 16; Length 642;
Best Local Similarity 99.5%; Pred. No. 8.5e-289;
Matches 639; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	19	MLRPGGLLLAVQLCTGTAVPSTGASKSRQAOQVQPSVAVSQSKPCYDNGKHQY	78
Db	1	MLRPGGLLLAVQLCTGTAVPSTGASKSRQAOQVQPSVAVSQSKPCYDNGKHQY	60
Qy	79	INQWERTYLGALVCTCYGSGRGNCEKPEABETCFDKYTGNTYRVGDTYERPKDSMI	138
Db	61	INQWERTYLGALVCTCYGSGRGNCEKPEABETCFDKYTGNTYRVGDTYERPKDSMI	120
Qy	139	WDCTCICAGRGRICTTANRCHGGQSKYKIGDTRRRPHETGGYMLCVCCLNGKGWTK	198
Db	121	WDCTCICAGRGRICTTANRCHGGQSKYKIGDTRRRPHETGGYMLCVCCLNGKGWTK	180

Qy 199 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNCNDQDTRTSY 258
Db 181 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNCNDQDTRTSY 240
Qy 259 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 318
Db 241 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 300
Qy 319 QPPYGHCVTDSDGVVYVSGMQLKTQGNKQMLCTCLNGVSCQETAVTQTYGGNSNGEPC 378
Db 301 QPPYGHCVTDSDGVVYVSGMQLKTQGNKQMLCTCLNGVSCQETAVTQTYGGNSNGEPC 360
Qy 379 VLPTYNDRDTSITSNVEODQKSFCTDHTVLVQTRGNSNGALCHFPFLYNNHNYTDC 438
Db 361 VLPTYNDRDTSITSNVEODQKSFCTDHTVLVQTRGNSNGALCHFPFLYNNHNYTDC 420
Qy 439 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQWDKQHDGMHMR 498
Db 421 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQWDKQHDGMHMR 480
Qy 499 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHRRHEBGMHLNCTCFGQGRGWKC 558
Db 481 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHRRHEBGMHLNCTCFGQGRGWKC 540
Qy 559 DPVQCODSETGTGYQIGDSWEKYVHGVRYQCYCGRGIGWCHCQPLQTYPSSSGPVEVF 618
Db 541 DPVQCODSETGTGYQIGDSWEKYVHGVRYQCYCGRGIGWCHCQPLQTYPSSSGPVEVF 600
Qy 619 ITETPSQNSHPIQWNAPOPSHISKYILRWPRVSIIPRNLGY 660
Db 601 ITETPSQNSHPIQWNAPOPSHISKYILRWPRVSIIPRNLGY 642

RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match 96.9%; Score 3664; DB 17; Length 642;
Best Local Similarity 99.5%; Pred. No. 8.5e-289;
Matches 639; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 19 MLRGPGLLLAVQCLGTAVPSTGASKSKRAQOMVQPSVAVSQSKPCYDNGKHQY 78
Db 1 MLRGPGLLLAVLCLGTAVPSTGASKSKRAQOMVQPSVAVSQSKPCYDNGKHQY 60
Qy 79 INQWERTYLGALVCTCYGSGRGNCSKPEAETCFDKYTGNTYRVDYTPKDSMI 138
Db 61 INQWERTYLGALVCTCYGSGRGNCSKPEAETCFDKYTGNTYRVDYTPKDSMI 120
Qy 139 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRRRPHETGGMLECYCLNGKGEWTC 198
Db 121 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRRRPHETGGMLECYCLNGKGEWTC 180
Qy 199 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNCNDQDTRTSY 258
Db 181 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNCNDQDTRTSY 240

Qy 259 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 318
Db 241 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 300
Qy 319 QPPYGHCVTDSDGVVYVSGMQLKTQGNKQMLCTCLNGVSCQETAVTQTYGGNSNGEPC 378
Db 301 QPPYGHCVTDSDGVVYVSGMQLKTQGNKQMLCTCLNGVSCQETAVTQTYGGNSNGEPC 360
Qy 379 VLPTYNDRDTSITSNVEODQKSFCTDHTVLVQTRGNSNGALCHFPFLYNNHNYTDC 438
Db 361 VLPTYNDRDTSITSNVEODQKSFCTDHTVLVQTRGNSNGALCHFPFLYNNHNYTDC 420
Qy 439 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQWDKQHDGMHMR 498
Db 421 SEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQWDKQHDGMHMR 480
Qy 499 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHRRHEBGMHLNCTCFGQGRGWKC 558
Db 481 CTCVNGRGWTCIAYSQLRDQCI VDDITVNVNDTFHRRHEBGMHLNCTCFGQGRGWKC 540
Qy 559 DPVQCODSETGTGYQIGDSWEKYVHGVRYQCYCGRGIGWCHCQPLQTYPSSSGPVEVF 618
Db 541 DPVQCODSETGTGYQIGDSWEKYVHGVRYQCYCGRGIGWCHCQPLQTYPSSSGPVEVF 600
Qy 619 ITETPSQNSHPIQWNAPOPSHISKYILRWPRVSIIPRNLGY 660
Db 601 ITETPSQNSHPIQWNAPOPSHISKYILRWPRVSIIPRNLGY 642

RESULT 3
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Query Match 96.4%; Score 3646.5; DB 16; Length 657;
Best Local Similarity 97.3%; Pred. No. 2.3e-287;
Matches 639; Conservative 0; Mismatches 3; Indels 15; Gaps 1;
Qy 19 MLRGPGLLLAVQCLGTAVPSTGASKSKRAQOMVQPSVAVSQSKPCYDNGKHQY 78
Db 1 MLRGPGLLLAVLCLGTAVPSTGASKSKRAQOMVQPSVAVSQSKPCYDNGKHQY 60
Qy 79 INQWERTYLGALVCTCYGSGRGNCSKPEAETCFDKYTGNTYRVDYTPKDSMI 138
Db 61 INQWERTYLGALVCTCYGSGRGNCSKPEAETCFDKYTGNTYRVDYTPKDSMI 120
Qy 139 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRRRPHETGGMLECYCLNGKGEWTC 198
Db 121 WDCTCIGAGRISCTIANRCHGGQSYKIGDTRRRPHETGGMLECYCLNGKGEWTC 180
Qy 199 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNCNDQDTRTSY 258
Db 181 PIAKCFDHAAGTSYVVGTEWKEPYQGMWVDTCTCLGSGRITCTSRNCNDQDTRTSY 240
Qy 259 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 318
Db 241 RIGDTSKKNRGNLLQICITGNRGWKCERHSTSVQTTSSGSPFFTDVRAAVYQPPHP 300
Qy 319 QPPYGHCVTDSDGVVYVSGMQLKTQGNKQMLCTCLNGVSCQETAVTQTYGGNSNGEPC 378

Result No.	Query			ID	Description	
	Score	Match	Length			
1	3595.5	95.1	2386	4	US-09-961-403-1	Sequence 1, Appl
2	3578.5	94.7	2386	2	US-09-016-366A-12	Sequence 12, Appl
3	3578.5	94.6	2446	2	US-08-551-356-2	Sequence 2, Appl
4	3578.5	94.6	2446	5	PCR-U893-12687-2	Sequence 2, Appl
5	3433.5	90.8	2327	6	5455158-1	Patent No. 5455158
6	3433.5	90.8	2327	6	5455158-1	Patent No. 5455158
7	3422.5	90.5	2324	1	US-08-283-857-1	Sequence 1, Appl
8	3422.5	90.5	2324	5	PCR-U895-09819-1	Sequence 1, Appl
9	3406.5	90.1	2231	1	US-08-153-799-16	Sequence 16, Appl
10	371	9.8	65	2	US-08-982-597A-19	Sequence 19, Appl
11	371	9.8	65	3	US-09-136-218-19	Sequence 19, Appl
12	298	7.9	663	4	US-09-194-468A-30	Sequence 30, Appl
13	297	7.9	631	3	US-08-448-489-17	Sequence 17, Appl
14	297	7.9	631	4	US-09-689-730-17	Sequence 17, Appl
15	297	7.9	660	3	US-08-704-711A-18	Sequence 18, Appl
16	297	7.9	660	3	US-09-521-220-18	Sequence 18, Appl
17	297	7.9	660	3	US-09-391-104-19	Sequence 19, Appl
18	297	7.9	660	4	US-09-917-254-89	Sequence 89, Appl
19	297	7.9	660	4	US-09-949-016-6512	Sequence 6512, Ap
20	297	7.9	660	4	US-09-949-016-7937	Sequence 7937, Ap
21	295	7.8	422	2	US-08-836-854-12	Sequence 12, Appl
22	294.5	7.8	188	1	US-08-142-449B-14	Sequence 14, Appl
23	294.5	7.8	474	2	US-08-836-854-9	Sequence 9, Appl
24	292.5	7.7	429	4	US-09-194-468A-45	Sequence 45, Appl
25	290	7.7	48	2	US-08-982-597A-22	Sequence 22, Appl
26	290	7.7	48	3	US-09-136-218-22	Sequence 22, Appl
27	285.5	7.6	707	3	US-08-704-711A-19	Sequence 19, Appl

Db 301 QPPYGHCVTDGSGVYVSGMQLKTCGNKQMLCTCLGNGVSCQETAFTVQTQTYGNSNGEPC 360
Qy 379 VLPFTYNDRT-----DSTTSNYEQDKYSFCTDHTVLVQTRGGNSGALC 423
Db 361 VLPFTYNGRTFYSCCTTEGRODGHLMWCSTTSNYEQDKYSFCTDHTVLVQTRGGNSGALC 420
Qy 424 HFPELYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 483
Db 421 HFPELYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 480
Qy 484 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDITVNVNDTFFKHEEGHM 543
Db 481 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDITVNVNDTFFKHEEGHM 540
Qy 544 LNCTCFQGRGWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCYGRGIGEWHCQ 603
Db 541 LNCTCFQGRGWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCYGRGIGEWHCQ 600
Qy 604 PLQTPSSSGPVEVFTETPSQNSHPIQWNAPOPSHISKYLWRP 650
Db 601 PLQTPSSSGPVEVFTETPSQNSHPIQWNAPOPSHISKYLWRP 647

RESULT 2

US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-016-366A-12

Query Match 94.7%; Score 3579.5; DB 2; Length 2386;
Best Local Similarity 96.8%; Pred. No. 1.7e-299;
Matches 626; Conservative 1; Mismatches 5; Indels 15; Gaps 1;

Qy 19 MLRGPGLILLVAVCLGTAVPSTGASKSRQQAQVQPOSQVAVSQSKFCYDNGKHQY 78
Db 1 MLRGPGLILLVAVCLGTAVPSTGASKSRQQAQVQPOSQVAVSQSKFCYDNGKHQY 60
Qy 79 INQWERTYLGNAVCTCYGSGRGFNCESPEABETCFDKYTGNTYRVGDTYERPKDSMI 138
Db 61 INQWERTYLGNAVCTCYGSGRGFNCESPEABETCFDKYTGNTYRVGDTYERPKDSMI 120
Qy 139 WDCTCIGAGRGISCTIANRCHGGGQYKIGDTRRRPHETGGYMLCEVCLGNGKGWETCK 198
Db 121 WDCTCIGAGRGISCTIANRCHGGGQYKIGDTRRRPHETGGYMLCEVCLGNGKGWETCK 180
Qy 199 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVMDCTCLGEGSGRITCTSRNCNDQDTRTSY 258
Db 181 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVMDCTCLGEGSGRITCTSRNCNDQDTRTSY 240
Qy 259 RIGDTSKDNRNGLQCICTGNGRGWKCERHTSVQTTSSGSGPFTDVRAAVYQOPHP 318
Db 241 RIGDTSKDNRNGLQCICTGNGRGWKCERHTSVQTTSSGSGPFTDVRAAVYQOPHP 300
Qy 319 QPPYGHCVTDGSGVYVSGMQLKTCGNKQMLCTCLGNGVSCQETAFTVQTQTYGNSNGEPC 378
Db 301 QPPYGHCVTDGSGVYVSGMQLKTCGNKQMLCTCLGNGVSCQETAFTVQTQTYGNSNGEPC 360
Qy 379 VLPFTYNDRT-----DSTTSNYEQDKYSFCTDHTVLVQTRGGNSGALC 423
Db 361 VLPFTYNGRTFYSCCTTEGRODGHLMWCSTTSNYEQDKYSFCTDHTVLVQTRGGNSGALC 420
Qy 424 HFPELYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 483
Db 421 HFPELYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 480
Qy 484 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDITVNVNDTFFKHEEGHM 543
Db 481 GDQWDKQHDGMHMRCTCVNGRGWETCIAYSQLRDQCI VDDITVNVNDTFFKHEEGHM 540
Qy 544 LNCTCFQGRGWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCYGRGIGEWHCQ 603
Db 541 LNCTCFQGRGWKCDPVDQCDSETGTFYIGDSWEKYVGVRYQCYCYGRGIGEWHCQ 600
Qy 604 PLQTPSSSGPVEVFTETPSQNSHPIQWNAPOPSHISKYLWRP 650
Db 601 PLQTPSSSGPVEVFTETPSQNSHPIQWNAPOPSHISKYLWRP 647

RESULT 3

US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Iranl, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551.356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 11, 2005, 07:55:44 ; Search time 244.157 Seconds
(without alignments)
1045.483 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781

Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILRWKRVSIIPRNLGY 660

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	3781	100.0	660	2	AAY28901 Human mig
2	3693.5	97.7	2182	8	ADR66462 Human pro
3	3693.5	97.7	2182	8	ADR66120 Human pro
4	3680.5	97.3	720	2	AAY28914 Fibronect
5	3680	97.3	642	8	ADR67316 Human bla
6	3680	97.3	642	8	ADSL17489 Amino aci
7	3680	97.3	642	8	ADR97658 Human fib
8	3664	96.9	642	8	ADQ39403 Human myo
9	3658	96.7	2447	4	AAM40434 Human pol
10	3658	96.7	2447	4	AAM40435 Human pol
11	3658	96.7	2447	4	AAM40433 Human pol
12	3658	96.7	2447	4	AAM40432 Human pol
13	3646.5	96.4	657	8	ADQ39409 Human myo
14	3599.5	95.2	2220	6	ABO01289 Human pro
15	3599.5	95.2	2266	6	ABR40124 Human cel
16	3599.5	95.2	2355	6	ABR58335 NM_00202
17	3599.5	95.2	2355	7	ADP65196 Human fib
18	3599.5	95.2	2355	8	ADG89560 Human fib
19	3599.5	95.2	2355	8	ADOS5175 Protein #
20	3599.5	95.2	2355	8	ADQ26085 Fibronect
21	3599.5	95.2	2355	8	ADQ29668 Human col
22	3599.5	95.2	2355	8	ADR67315 Human bla
23	3599.5	95.2	2386	8	ADOS5174 Protein #
24	3599.5	95.2	2474	4	ABG22279 Novel hum
25	3599.5	95.1	2386	5	AAO17353 Human fib

26	3595.5	95.1	2386	6	ABR81866 Human fib
27	3595.5	95.1	2386	7	ADD18770 Human dis
28	3595.5	95.1	2386	7	ADG63324 Human pro
29	3595.5	95.1	2386	8	ADR99200 Fibronect
30	3593.5	95.0	2446	3	ABR50377 Human fib
31	3592.5	95.0	984	8	ADQ39406 Human myo
32	3592.5	95.0	2265	4	AAM38647 Human pol
33	3592.5	95.0	2296	8	ADQ39412 Human myo
34	3592.5	95.0	2330	4	AAM38646 Human pol
35	3592.5	95.0	2355	4	AAM38649 Human pol
36	3592.5	95.0	2355	8	ADQ39415 Human myo
37	3592.5	95.0	2355	8	ADQ39404 Human myo
38	3592.5	95.0	2386	4	AAM38648 Human pol
39	3592.5	95.0	2386	8	ADQ39408 Human myo
40	3586.5	94.9	2477	2	AAW99595 Human fib
41	3584.5	94.8	2220	8	ADN95950 Human NOV
42	3579.5	94.7	2386	2	AAW63171 Amino aci
43	3578.5	94.6	2446	2	AAR60021 Fibronoge
44	3546	93.8	2352	8	ADK00410 Fibronect
45	3448.5	91.2	2328	4	AAG68182 Fibronect

ALIGNMENTS

RESULT 1

AAY28901

ID AAY28901 standard; protein; 660 AA.

AC AAY28901;

XX 21-SEP-1999 (first entry)

DE Human migration stimulating factor (MSF) 1-alpha protein.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;

KW wound healing; scarring; MSF1-alpha.

XX Homo sapiens.

XX WO9931233-A1.

PD 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

XX 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

DR WPI; 1999-430039/36.

DR N-PSDB; AAX81299.

XX Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX Claim 1; Page 53; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. The present sequence represents the human MSF1-alpha protein

XX Sequence 660 AA;

XX Query Match 100.0%; Score 3781; DB 2; Length 660;

XX Best Local Similarity 100.0%; Pred. No. 7.1e-242; Indels 0; Gaps 0;

XX Matches 660; Conservative 0; Mismatches 0;

QY 1 NLVATCLPVRASLPHRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 60
 Db 1 NLVATCLPVRASLPHRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 60
 QY 61 VAVSQKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCESKPEAEETCFDKYT 120
 Db 61 VAVSQKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCESKPEAEETCFDKYT 120
 QY 121 GNTYRVGDTYRDKSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGG 180
 Db 121 GNTYRVGDTYRDKSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGG 180
 QY 181 YMLECVCLGNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 240
 Db 181 YMLECVCLGNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 240
 QY 241 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 300
 Db 241 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 300
 QY 301 SGPTFDVRAAVYQPPHPPQPPYGHCVTDSGVVSVGNQWLKTQGNKQMLCTCLGNGVSC 360
 Db 301 SGPTFDVRAAVYQPPHPPQPPYGHCVTDSGVVSVGNQWLKTQGNKQMLCTCLGNGVSC 360
 QY 361 QETAVTQYVGNNGEPCVLPFTYNDRTDSTTSNVEQDKYSFCTDHTVLVQTRGNSNG 420
 Db 361 QETAVTQYVGNNGEPCVLPFTYNDRTDSTTSNVEQDKYSFCTDHTVLVQTRGNSNG 420
 QY 421 ALCHFPFLYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEBICITNEGVM 480
 Db 421 ALCHFPFLYNNHNTDCTSEGRDNMKWCGTTQNYDADQKFGFCPMAAHEBICITNEGVM 480
 QY 481 YRIGQDQKQDGMHMRCTCVGNRGEWTCIAYSQLRDQICVDITVNDVTFKRHEE 540
 Db 481 YRIGQDQKQDGMHMRCTCVGNRGEWTCIAYSQLRDQICVDITVNDVTFKRHEE 540
 QY 541 GHMLNCTCFGQGRGWKCDPVDQCODSETGTYQIGDSWEKYVHGVYQCYGGRIGEW 600
 Db 541 GHMLNCTCFGQGRGWKCDPVDQCODSETGTYQIGDSWEKYVHGVYQCYGGRIGEW 600
 QY 601 HCQPLQTYPSSSGPVEVITETPSQPNHPIQWNAFPQPSHISKYILRWPRVSIIPRNLGY 660
 Db 601 HCQPLQTYPSSSGPVEVITETPSQPNHPIQWNAFPQPSHISKYILRWPRVSIIPRNLGY 660

RESULT 2

AD66462
 ID AD66462 standard; protein; 2182 AA.

XX AC AD66462;
 XX DT 02-DEC-2004 (first entry)
 XX DE Human prostatic carcinoma derived protein SEQ ID 316 #2.

XX KW human; cytostatic; diagnosis; prostatic cancer;
 XX KW differential expression analysis.
 XX OS Homo sapiens.

XX PN WO2004076614-A2.
 XX PD 10-SEP-2004.
 XX PF 22-FEB-2004; 2004WO-DE000433.
 XX PR 27-FEB-2003; 2003DE-01009985.
 XX PR 14-MAY-2003; 2003DE-01022134.

XX (HINZ/) HINZMANN B.
 PA (DAHL/) DAHL E.
 PA (ROSE/) ROSENTHAL A.
 PA (HERM/) HERMANN K.

(PILA/) PILARSKY C.

Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarczyk C, Specht T;
 Schmitt A, Beckmann G, Bruemendorf T, Kimmann H, Roepcke S;
 Xinzhang L, Staub E;
 WPI; 2004-653386/63.

New nucleic acids, and encoded proteins, from prostatic cancer tissue,
 useful for diagnosis, treatment and in screening for specific binding
 agents.

Claim 2; Page 1298; 1607pp; German.

This invention describes novel cytosolic polynucleotide and polypeptide
 sequences which can be used in a method for diagnosing prostatic cancer
 or the risk of developing prostatic cancer. Diagnosis is based on
 determining over transcription or over expression of the sequences in
 prostatic tissue. Screening for inhibitors of the sequences or detection
 substances involves a binding assay, any compounds that bind are
 selected, optionally after deconvolution of mixtures. Detection of a
 predetermined minimum level of the reporter indicates the presence of
 tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
 short-interfering RNA or ribozymes; an organic molecule of molecular
 weight below 5000, preferably 300, that binds to the polypeptide; an
 aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
 polypeptide, preferably humanised or human; an anti-idiotypic, non-human
 (monoclonal) antibody directed against Ab or any of the above derivatised
 with a reporter group, cell toxin, immunostimulatory molecules and/or
 radioisotope. The polynucleotides are identified in human prostatic
 cancer by differential expression analysis, using DNA microarrays,
 between normal and tumorous tissues, with (over)expression being detected
 by quantitative PCR. Analysis of prostatic cancer samples showed that
 CD24 was upregulated in many of them. Sections of tissue, isolated from
 prostatic cancer patients, or subjects at risk, were incubated
 sequentially with anti-human CD4 murine monoclonal antibodies;
 biotinylated second antibody; streptavidin-conjugated horseradish
 peroxidase and then diaminobenzidine as colour former (brown). The
 samples were counterstained with hemalum (blue). Malignant cells stained
 strongly but non-malignant cells only weakly. In 15 of 63 samples of
 adenocarcinoma, membrane and cytoplasmic staining was very strong, and
 lymph node metastases were also stained. AD66462-AD66494 represent the
 polynucleotide and polypeptide sequences used in the method of the
 invention.

Sequence 2182 AA;

Query Match 97.7%; Score 3693.5; DB 8; Length 2182;
 Best Local Similarity 97.4%; Pred. No. 1.7e-235;
 Matches 648; Conservative 0; Mismatches 2; Indels 15; Gaps 1;
 QY 1 NLVATCLPVRASLPHRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 60
 Db 71 NLVATCLPVRASLPHRLNMLRPGPGLLLAVQCLGTAVPSTGASKSKRQAOQVQPOSP 130
 QY 61 VAVSQKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCESKPEAEETCFDKYT 120
 Db 131 VAVSQKPGCYDNGKHQYQINQWERTYLGNALVCTCYGSGRGFNCESKPEAEETCFDKYT 190
 QY 121 GNTYRVGDTYRDKSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGG 180
 Db 191 GNTYRVGDTYRDKSMIWDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGG 250
 QY 181 YMLECVCLGNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 240
 Db 251 YMLECVCLGNGKGWETCKPIAEKCFDHAAGTSYVVGWETWEKPYQGMWMDCTCLGEGSGR 310
 QY 241 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 300
 Db 311 ITCTSRNRNCNDQDTRTSYRIGDTSKDNKGNLLQICITGNRGEWKCEHRTSVQTTSSG 370
 QY 301 SGPTFDVRAAVYQPPHPPQPPYGHCVTDSGVVSVGNQWLKTQGNKQMLCTCLGNGVSC 360

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:46:42 ; Search time 13726.3 Seconds
(without alignments)
1830.243 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781
Sequence: 1 NLVATCPVRASLPHRLML.....ISKYILWRPVSIPPRNLGY 660

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_b/US09581651/runat_07112005_092223_28789/app_query.fasta_1.1834
-DB=EST -QFMT=fascap -SUPPIX=p2n.rst -MINWATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651 @CIGN 1 1 8180 @runat_07112005_092223_28789 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3685.5	97.5	7885	3 CR749281	CR749281 Homo sapi
2	3682.5	97.4	7777	3 CR749316	CR749316 Homo sapi
3	3661.5	96.8	8121	3 CR749317	CR749317 Homo sapi
4	3658.5	96.8	7501	3 BC078656	BC078656 Homo sapi
5	3419.5	90.4	8315	3 AK090135	AK090135 Mus muscu
6	3308	87.5	8329	3 AK090130	AK090130 Mus muscu
7	1573.5	41.6	943	5 BX391752	BX391752 BX391752
8	1564	41.4	1044	5 BX398837	BX398837 BX398837
9	1526	40.4	2107	3 AK054456	AK054456 Mus muscu

10	1490.5	39.4	836	7	CN538822	UI-M-HS0-
11	1480	39.1	1052	5	BX380582	BX380582
12	1478	39.1	816	1	AU141008	AU141008
13	1464	38.7	859	7	CF616056	AGENCOURT
14	1460	38.6	766	7	CN419594	170004345
15	1451.5	38.4	861	1	AU140971	AU140971
16	1450	38.3	1026	5	BX386270	BX386270
17	1419.5	37.5	781	7	CN534124	UI-M-HOO-
18	1415	37.4	765	4	BM715855	UI-E-EJO-
19	1402.5	37.1	1050	5	BX417945	BX417945
20	1400	37.0	808	7	CK638401	UI-M-HOO-
21	1390	36.8	837	1	AU140973	AU140973
22	1386	36.7	731	1	AU140526	AU140526
23	1381	36.5	737	1	AU140993	AU140993
24	1369	36.2	959	5	BU109952	603127860
25	1356	35.9	739	1	AL706215	DKFZp686J
26	1351	35.7	695	7	CN419477	170004247
27	1349	35.7	730	1	AU140834	AU140834
28	1340	35.4	772	7	CN530412	UI-M-HOO-
29	1329.5	35.2	781	7	CN161442	950735 MA
30	1323	35.0	699	1	AU140889	AU140889
31	1308	34.6	723	7	CN419622	170005319
32	1304.5	34.5	780	7	CN162843	952271 MA
33	1303	34.5	738	6	CD613781	55110394J
34	1302.5	34.4	748	1	AU140948	AU140948
35	1298	34.3	667	1	AU140910	AU140910
36	1295.5	34.3	796	1	AU140814	AU140814
37	1291.5	34.2	748	1	AU140991	AU140991
38	1289	34.1	676	1	AL603362	DKFZp686C
39	1284	34.0	705	1	AU140789	AU140789
40	1283.5	33.9	758	1	AU140556	AU140556
41	1282.5	33.9	941	5	BQ928772	AGENCOURT
42	1281	33.9	734	1	AU140802	AU140802
43	1279.5	33.8	851	1	AU141024	AU141024
44	1271	33.6	673	1	AU140450	AU140450
45	1245.5	32.9	759	1	AU140735	AU140735

ALIGNMENTS

RESULT 1

CR749281 LOCUS CR749281 7885 bp mRNA linear HTC 19-AUG-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp686F10164 (from clone DKFZp686F10164).
ACCESSION CR749281
VERSION CR749281.1 GI:51476291
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 7885)
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobbo, G., Han, M. and Wiemann, S.
The German cDNA Consortium
CONSTRM Direct Submission
TITLE Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764
JOURNAL Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp686F10164) is available at the R2PD Deutsches
Reasourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact R2PD for ordering:
http://www.r2pd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686F10164
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source
1..7885
/organism="Homo sapiens"
/mol_type="mRNA"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 15:20:41 ; Search time 3388.86 Seconds
(without alignments)
1610.594 Million cell updates/sec

Title: US-09-581-651D-1
Perfect score: 3781
Sequence: 1 NLVATCLFVRASLPHRLNML.....ISKYILWRPVSIPPRNLGY 660

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9794790 seqs, 4134909567 residues

Total number of hits satisfying chosen parameters: 19589580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US09581651/runat_07112005_092226_28867/app_query.fasta_1.1834
-DB=PublishedApplications_NA_QPMT=fastp -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCUI=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl6sum62
-TRANS=human40.cdi -LIST=45 -DALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -WMAP=US09581651 @CGN 1 1 1041 @runat_07112005_092226_28867
-NCPUI=6 -ICPU=3 -NO_MMAP -LARGESQURY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRADs=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DEUEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
27: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
28: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3765	99.6	2127	22	US-10-956-157-4288 Sequence 4288, Ap
2	3759	99.4	2443	20	US-10-741-601-70 Sequence 70, Appl
3	3759	99.4	2443	22	US-10-741-600-238 Sequence 238, App
4	3758	99.4	2127	17	US-10-210-120-49 Sequence 49, Appl
5	3758	99.4	2127	24	US-10-909-035-49 Sequence 49, Appl
6	3741.5	99.0	2488	20	US-10-741-601-75 Sequence 75, Appl
7	3741.5	99.0	2488	22	US-10-741-600-244 Sequence 244, App
8	3693.5	97.7	8027	18	US-10-447-161-8 Sequence 8, Appl
9	3693.5	97.7	8027	20	US-10-734-564-27 Sequence 27, Appl
10	3693.5	97.7	8027	22	US-10-852-335A-53 Sequence 53, Appl
11	3693.5	97.7	8027	24	US-10-287-436A-81 Sequence 81, Appl
12	3693.5	97.7	8815	22	US-10-868-577A-62 Sequence 62, Appl
13	3693.5	97.7	8815	20	US-10-868-549-21 Sequence 21, Appl
14	3687.5	97.5	6510	22	US-10-741-601-72 Sequence 72, Appl
15	3687.5	97.5	6510	20	US-10-741-600-241 Sequence 241, App
16	3687.5	97.5	7823	20	US-10-741-601-77 Sequence 77, Appl
17	3687.5	97.5	7823	22	US-10-741-600-245 Sequence 245, App
18	3687.5	97.5	7848	20	US-10-741-601-78 Sequence 78, Appl
19	3687.5	97.5	7848	22	US-10-741-600-246 Sequence 246, App
20	3687.5	97.5	7935	20	US-10-741-601-74 Sequence 74, Appl
21	3687.5	97.5	7935	22	US-10-741-600-240 Sequence 240, App
22	3687.5	97.5	7959	20	US-10-741-601-81 Sequence 81, Appl
23	3687.5	97.5	7959	22	US-10-741-600-249 Sequence 249, App
24	3687.5	97.5	8013	20	US-10-741-601-71 Sequence 71, Appl
25	3687.5	97.5	8013	22	US-10-741-600-242 Sequence 242, App
26	3687.5	97.5	8155	20	US-10-741-601-79 Sequence 79, Appl
27	3687.5	97.5	8155	22	US-10-741-600-247 Sequence 247, App
28	3687.5	97.5	8226	20	US-10-741-601-69 Sequence 69, Appl
29	3687.5	97.5	8226	22	US-10-741-600-237 Sequence 237, App
30	3687.5	97.5	8278	20	US-10-741-601-82 Sequence 82, Appl
31	3687.5	97.5	8278	22	US-10-741-600-250 Sequence 250, App
32	3687.5	97.5	8332	20	US-10-741-601-73 Sequence 73, Appl
33	3687.5	97.5	8332	22	US-10-741-600-239 Sequence 239, App
34	3687.5	97.5	8371	20	US-10-741-601-76 Sequence 76, Appl
35	3687.5	97.5	8371	22	US-10-741-600-243 Sequence 243, App
36	3686.5	97.5	7795	16	US-10-084-817-2 Sequence 2, Appl
37	3686.5	97.5	7867	14	US-10-098-841-6 Sequence 6, Appl
38	3686.5	97.5	8044	17	US-10-240-965-121 Sequence 121, App
39	3686.5	97.5	8044	24	US-10-765-700-135 Sequence 135, App
40	3686.5	97.5	8062	14	US-10-098-841-5 Sequence 5, Appl
41	3686.5	97.5	8137	14	US-10-098-841-8 Sequence 8, Appl
42	3686.5	97.5	8230	14	US-10-098-841-7 Sequence 7, Appl
43	3605.5	95.4	7361	19	US-10-236-392-3 Sequence 3, Appl
44	3547.5	93.8	8216	24	US-10-450-763-22270 Sequence 22270, A
45	3527	93.3	4295	18	US-10-144-194A-51 Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-10-956-157-4288
; Sequence 4288, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCES: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956.157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: Patent version 3.2

```
; SEQ ID NO 4288
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-4288

Alignment Scores:
Pred. No.: 0 Length: 2127
Score: 3765.00 Matches: 658
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 2
Query Match: 99.58% Indels: 0
DB: 22 Gaps: 0

US-09-581-651D-1 (1-660) x US-10-956-157-4288 (1-2127)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 3 AACTTGGTGGCAACTTGGCTCCCGGTGGGGGCTCTCTCCCCACCGTCTCAACATGCTT 62
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 63 AGGGTCCGGGGCCGGGCTGCTGCTGCTGCTGCCGCTCCAGTGCCTGGGGACAGCGGTGCC 122
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
Db 123 TCCACGGGAGGCTCGAAGCAAGAGGAGGCTCAGCAATGGTTTCAGCCCGAGTCCCGG 182
Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
Db 183 GTGGCTGTCAAGCAAGCCCGTGTGTATGACATGGAACACTATCAGATAAT 242
Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
Db 243 CAACAGTGGGAGCGGACCTACCTAGGCAATGGTGTGTGTACTTGTATGGAGGAGC 302
Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluThrCysPheAspLysTyrThr 120
Db 303 CGAGGTTTTAACCTGCGAGAGTAACCTGAAGCTGAAGAGACTTGTCTTGAACAAGTACACT 362
Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 363 GGGAACTATACCGAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGAC 422
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgLysSerCysThrIleAlaAsnArgCysHis 160
Db 423 TGTACTCTGCATCGGGGCTGGGCGAGGAGAAATAAGCTGTACCATCGCAAAACCGCTGCCAT 482
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTyrArgArgProHisGluThrGlyGly 180
Db 483 GAAGGGGTGAGTCTCTACAAGATTGGTGACACCTGGAGAGACCATGAGACTGGTGGT 542
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle 200
Db 543 TACATGTTAGAGTGTGTGCTTCTGTAATGGAAGAGAGAAATGACCTGCAAGCCCAT 602
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db 603 GCTGAGAAAGTGTGTTTGTATCATCTCTGGGACTTCTCTATGTGTGGGAGAAACGTGGGAG 662
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 663 AAGCCCTTACCAGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 723 ATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACACAAAGGACATCCTATAGAATT 782
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 783 GGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCAACAGGC 842
Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
843 AAGCGCGAGGAGTGGAGTGTGAGAGGCGACACCTCTGTGAGACCAATCATCGAGCGGA 902
301 SerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisProGlnPro 320
903 TCTGGCCCTTCCACCGATGTTCTGTCAGCTGTTTACCAACCGAGCGCTCCACCCAGCCT 962
321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
963 CTCTCCCTATGGCCACTGTGTGCAGACAGTGGTGTGTGTCTCTCTGTGGGATGCACTGG 1022
341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
1023 CTGAAGACACAGAGAAATAAGCAATGCTTTTGCAGCTGCTGGGCAACGGAGTCAGCTGC 1082
361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
1083 CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGTCTTA 1142
381 ProPheThrTyrAsnAspArgThrAspSerThrThrSerAsnTyrGluGlnAspGlnLys 400
1143 CCATTACCTTACAAATGGCAGAGCTGTGCAGCACAACCTTCGAATATATGAGCAGGACGAGAA 1202
401 TyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGly 420
1203 TACTCTTCTGCACAGACACACTGTTTTTGGTTTCAGACTCGAGGAGGAAATTCCAATGGT 1262
421 AlaLeuCysHisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGlu 440
1263 GCCTTGTGCCACTTCCCTTCTCTATACAAACACCAATTACACTGATTCGACTTCTGAG 1322
441 GlyArgArgAspAsnMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLys 460
1323 GGCAGAGAGACAACTGAAGTGGTGTGGGACCAACAGAACTATGATGCCGACACAGAG 1382
461 PheGlyPheCysProMetAlaAlaHisGluGluLeuCysThrThrAsnGluGlyValMet 480
1383 TTTGGGTTCTGCCCTTCGCTGCCACAGAGAAATCTGCACAACCAATGAAGGGTCTATG 1442
481 TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThr 500
1443 TACCGCATTTGAGATCATGATGGGATTAAGCAGCATGACATGGGTTCACATGATGAGTGCACG 1502
501 CysValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGln 520
1503 TGTGTTGGGAATGTTCTGGGAAATGACATGCAATGCTTACTTCCGACTTCGAGATCAG 1562
521 CysIleValAspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGlu 540
1563 TGCATTGTTGATGACATCACTTACAATGTGAACGACACATTTCCACAAAGCGTCATGAAGAG 1622
541 GlyHisMetLeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspPro 560
1623 GGGCACATGCTGAACTGTACATGCTTCCGTTCAGGCTCGGGCAGGTGGAGTGTGATCCC 1682
561 ValAspGlnCysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGlu 580
1683 GTCCAGCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATGGAGATTCATGGAG 1742
581 LysTyrValHisGlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrp 600
1743 AAGTATGTGATGCTGTGCAGATACAGTGTCTTCTGCTATGCGCTGTCATGGGAGTGG 1802
601 HisCysGlnProLeuGlnThrTyrProSerSerSerGlyProValGluValPheIleThr 620
1803 CATTGCCAACTTACAGACCTTATCAAGCTCAAGCTGGTCTGTGCAAGATATTATTAAT 1862
621 GluThrProSerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHis 640
1863 GAGACTCGAGTCAAGCCCACTCCCAACCCATCCAGTGGATGCAACCAACCAACCACTCTAC 1922
641 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 660
```

Db 1803 CATTGCCAACCTTTACAGACCTATCCAGGCTCAAGTGGTCTCTGCGAAGTATTTTACACT 1862
 Qy 621 GluThrProSerGlnProAsnSerHisProIleGlnTrpAenAlaProGlnProSerHis 640
 Db 1863 GAGACTCGAGTCAAGCCAACTCCACCCCATCCAGTGGATGACACACAGCCATCTCAC 1922
 Qy 641 IleserlystYrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 660
 Db 1923 ATTTCCAAGTACATTTCTCAGGTGAGACCTGTGAGTATCCACCCAGAAACCTTGGATAC 1982

RESULT 2

ADR67201

ID ADR67201 standard; DNA; 2358 BP.

XX AC ADR67201;

DT 02-DEC-2004 (first entry)

XX DE Human bladder cancer associated nucleotide sequence.

XX KW bladder cancer tissue; bladder cancer; cytostatic; gene; ds.

XX OS Homo sapiens.

XX PN WO2004076613-A2.

XX PD 10-SEP-2004.

XX PF 24-FEB-2004; 2004WO-DE000364.

XX PR 26-FEB-2003; 2003DE-01009729.

XX PA (HERR/) HERR A.

XX PA (HINZ/) HINZMANN B.

XX PA (DAHL/) DAHL E.

XX PA (STAU/) STAU E.

XX PA (PILA/) PILARSKY C.

XX PA (SPEC/) SPECHT T.

XX PI Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;

XX WPI; 2004-653385/63.

DR New nucleic acids, and encoded proteins, from bladder cancer tissue,
 PT useful for diagnosis, treatment and in screening for specific binding
 PT agents.

PS Claim 1; Fig 3; 112pp; German.

XX The present invention describes nucleic acids (I) associated with bladder
 CC cancer tissue. Also described: (1) peptides and proteins (II) containing
 CC an amino acid sequence encoded by (I); (2) a method for diagnosing
 CC bladder cancer (BC), or monitoring its progression, that uses (I), (II)
 CC or agents (Z) that inhibit, or bind to, (I) or (II), labelled with a
 CC reporter; and (3) a method for treating BC that uses (I), (II) or (Z).
 CC (I) and (II) have cytostatic activity. (I) and (II) can be used to detect
 CC (and monitor progression of) bladder cancer (BC), or the risk of
 CC developing it; to screen for specific binding agents (Z), and to treat
 CC BC. (Z) are also useful as diagnostic and therapeutic agents. The present
 CC sequence represents a human nucleotide sequence associated with bladder
 CC cancer, which is used in the exemplification of the present invention.

SQ Sequence 2358 BP; 611 A; 587 C; 639 G; 521 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.05e-245 Length: 2358
 Score: 3774.00 Matches: 659
 Percent Similarity: 99.85% Conservative: 0
 Best Local Similarity: 99.85% Mismatches: 1
 Query Match: 99.81% Indels: 0
 Gaps: 13

DB:

US-09-581-651D-1 (1-660) x ADR67201 (1-2358)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
 Db 214 AACTTGGTGGCAACTTGCCTCCCGGTGGCGGTCTCTCCCCACCGCTCTCAACATGCTT 273
 Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
 Db 274 AGGGGTCCGGGGCCGGGCTGCTGCTGCGCGTCCAGTGCCTGGGGACAGCGGTGCC 333
 Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
 Db 334 TCCACGGGAGCCTCGAAGACCAAGAGCAGGCTCAGCAATGGTTTCAGCCCGCAGTCCCG 393
 Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
 Db 394 GTGGCTGTCAAGCAAGCCCGGTTGTTATGACAAATGGAACAACACTATCAGATAAT 453
 Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlySer 100
 Db 454 CAACAGTGGAGCGGACCTACCTAGGCATGCGTGGTGGTTGTACTTGTATTATGGAGGAGC 513
 Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120
 Db 514 CGAGGTTTTAACTCGCGAGAGTAACTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACT 573
 Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
 Db 574 GGGAAACACTTACCAGTGGGTGGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGAC 633
 Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
 Db 634 TGTACCTGCATCGGGCTGGGCGAGGAGAAATAGCTGTACCATCGAACAACCGCTGCAT 693
 Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
 Db 694 GAAGGGGTCAGTCTCTACAAGATTGGTGACACCTGGAGGAGACCACTAGACTGTGGT 753
 Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle 200
 Db 754 TACATGTTAGAGTGTGTGTCTTGGTAAATGGAAGAGAGAGATGGACCTGCAAGCCCAT 813
 Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
 Db 814 GCTGAGAAGTGTTTTGATCATGCTGCTGGGACTTCTTATGTTGGTGGAGAACCTGGAG 873
 Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
 Db 874 AAGCCCTACCAGGCTGGATGATGGTAGATTGTACTTGGCTGGAGAAAGGAGCGGACGC 933
 Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgile 260
 Db 934 ATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTATAGAATT 993
 Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
 Db 994 GGAGACACCTCGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1053
 Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
 Db 1054 AAGCGCGGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGACAGACCACTCAGGCGGA 1113
 Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
 Db 1114 TCTGGCCCTTTCACCGATGTTCTGTCAGCTGTTTACCACCGCAGCCTCACCCCGCCCT 1173
 Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
 Db 1174 CTTCCCTATGGCCACTGTGTGCACAGACAGTGGTGGTGTGTCTACTCTGTGGGATGCAGTGG 1233
 Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
 Db 1234 CTGAAGACACACAGGAAATAAGCAATGCTTTGACAGTGCCTTGGCAACCGAGTGCAGTGC 1293

```
Qy 361 GlnGluThrAlaValThrGlnThrTyGlyGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1294 CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTCTTTA 1353
Qy 381 ProPheThrTyAsnAspArgThrAspSerThrThrSerAsnTyrGluGlnAspGlnlys 400
Db 1354 CCATTACCTCAATGGCAGACGACAGCACAACCTTCGAATTTATGACGAGCACAGAA 1413
Qy 401 TyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGly 420
Db 1414 TACTCTTCTGCACAGACCACTGTTTGGTTTCAGACTCGAGGAGGAAATCCAAATGGT 1473
Qy 421 AlaLeuCysHisLePheProPheLeuTyRanAsnHisAsnTyrThrAspCysThrSerGlu 440
Db 1474 GCCTTGTGCCACTTCCCTTCCATACAAACACCAATATCACTGATTCACATTCCTGAG 1533
Qy 441 GlyArgArgAspAsnMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnlys 460
Db 1534 GGCAGAGAGACACATGAGTGGTGGGACCACAGAACTATGATGCCGACCAAGAG 1593
Qy 461 PheGlyPheCysProMetAlaAlaHisGluGluLeCysThrThrAsnGluGlyValMet 480
Db 1594 TTTGGGTTCTGCCCATGGCTGCCACGAGGAAATCTGCACAACCAATGAAGGGGTCTATG 1653
Qy 481 TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThr 500
Db 1654 TACCAGATTGGAGATCACTGGGATGAAGCAGCATGACATGGGTCAATGATGAGTGACG 1713
Qy 501 CysValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGln 520
Db 1714 TGTGTTGGGAATGGTCTGGGGAATGGACATGCTACTGCTACTCGACCTTCGAGATCAG 1773
Qy 521 CysIleValAspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGlu 540
Db 1774 TGCATTGTGATGACATCACTTACAAATGTGAACGACACATTCACAACGCGTCATGAAGAG 1833
Qy 541 GlyHisMetLeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspPro 560
Db 1834 GGGCAGATGCTGAACCTGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAATGTGATCCC 1893
Qy 561 ValAspGlnCysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGlu 580
Db 1894 GTCGACCAATGCCAGGATTCAGAGCTGGGAGCTTTTATCAAAATGGAGATTCATGGGAG 1953
Qy 581 LysTyrValHisGlyValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrp 600
Db 1954 AAGTATGTGATGGTGTGCAGATACCAGTGTCTACTGCTATGGCCGTGGCATTTGGGAGTGG 2013
Qy 601 HisCysGlnProLeuGlnThrTyrProSerSerSerGlyProValGluValPheIleThr 620
Db 2014 CATTCGCCAACCTTTACAGACCTTATCCAAAGCTCAAGTGGTCTGTGCAAGTATTTATCACT 2073
Qy 621 GluThrProSerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHis 640
Db 2074 GAGACTCCAGTCCAGCCCACTCCACCCCACTCCAGTGGATGACCAACGACCATCTAC 2133
Qy 641 IleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeuGlyTyr 660
Db 2134 ATTTCCAAAGTACATTTCTCAGTGGAGACCTGTGTGAGTATCCCAACCCAGGAAACCTTGGATAC 2193
RESULT 3
ADQ38575
ID ADQ38575 standard; DNA; 2443 BP.
XX
AC ADQ38575;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human SNP containing myocardial infarction-associated gene, SEQ ID 238.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiac; gene therapy; human; gene; ds.
```

```
XX Homo sapiens.
OS
FN WO2004058052-A2.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
XX P-PSDB; ADQ39403.
XX
PT Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
PS Claim 7; SEQ ID NO 238; 145pp; English.
XX
CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiac activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
SQ Sequence 2443 BP; 599 A; 628 C; 682 G; 532 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 1,13e-244 Length: 2443
Score: 3759.00 Matches: 657
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 3
Query Match: 99.42% Indels: 0
DB: 13 Gaps: 0

US-09-581-651D-1 (1-660) x ADQ38575 (1-2443)
Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 319 AACTTGGTGGCACTTGGCTCCCGTGGGGGCTCTCTCCCCACCGCTCAACATGCTT 378
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
|||||
```


PD 13-FEB-2003.
XX
XX 02-AUG-2002; 2002WO-US024567.
PF
XX 02-AUG-2001; 2001US-0309581P.
PR 15-NOV-2001; 2001US-0334468P.
PR 01-AUG-2002; 2002US-00210120.
XX
XX (UNMI) UNIV MICHIGAN.
PA
XX
XX Rubin MA, Chinnaiyan AM, Sreekumar A;
PI WPI; 2003-278396/27.
XX
XX Characterizing prostate tissue comprises providing a prostate tissue
PT sample from a subject and detecting the presence or absence of expression
PT of hepsin, pim-1 or EZH2.
XX
XX Disclosure; SEQ ID NO 49; 297pp; English.
XX
XX This invention relates to a novel method of characterising prostate
CC tissue in a subject and to compositions and methods for cancer
CC diagnostics, including cancer markers, in particular prostate cancer.
CC Prostate cancer (PCA) is a leading cause of male cancer-related death.
CC Additional serum and tissue biomarkers would aid diagnosis. The invention
CC may provide means of producing compounds with a cytostatic activity or
CC allow the development of gene therapy. The methods of the invention
CC useful for characterising prostate tissue in a subject, screening
CC compounds, characterising inconclusive prostate biopsy tissue in a
CC subject, detecting AMACR (alpha-methylacyl-coenzyme A racemase)
CC expression in a bodily fluid, characterising tissue in a subject,
CC diagnosing cancer in a subject and inhibiting the growth of cells. The
CC present sequence is a DNA sequence which is preferably utilised in the
CC method of the invention.
XX
XX SQ Sequence 2127 BP; 552 A; 511 C; 576 G; 488 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,126-244 Length: 2127
Score: 3758.00 Matches: 657
Percent Similarity: 99.55% Conservative: 0
Best Local Similarity: 99.55% Mismatches: 3
Query Match: 99.39% Indels: 0
DB: 10 Gaps: 0

US-09-581-651D-1 (1-660) x ADD18477 (1-2127)

QY 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
DB 3 AACTTGGTGGCAACTTGCCTCCGGTGGGGGCTCTCCCCCACCCTCAACATGCTT 62
QY 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
DB 63 AGGGGTCCGGGGCCGGGGTGTGCTGTGTCGGCGCTCTGTGCTGGGACAGCGGTGCC 122
QY 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
DB 123 TCCAGGGAGCCCTCGAAGACGAAGAGGAGGCTGAGCAATGGTTACGCCCCAGTCCCG 182
QY 61 ValAlaValSerGlnSerLysProGlyCysThrAspAsnGlyLysHisThrGlnLeuAsn 80
DB 183 GTGGCTGTGAGTCAAGCAAGCCCGTGTGTTATGACATGGAACACACTATCAGATAAAT 242
QY 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlySer 100
DB 243 CAACAGTGGAGCGGACCTACCTAGGCAATGGCTTGGTTGTACTTGTATGGAGGAAGC 302
QY 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluThrCysPheAspLysTyrThr 120
DB 303 CGAGGGTTTAACTGCGAGAGTAACTGAGCTGGAAGACCTTGTGACAAAGTACACT 362
QY 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140

DB 363 GGAAACACTTACCAGTGGGTGACACTTATAGCGTCCTAAAGACTCCATGATCTGGAC 422
QY 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
DB 423 TGTACCTGCTGCTGGGGCTGGCGAGGAGAGTAAGCTGTACCATCGCAACCCGCTGCAT 482
QY 161 GluGlyGlyGlnSerTyrIleGlyAspThrTyrArgArgProHisGluThrGlyGly 180
DB 483 GAAGGGGTCACTCTACAAGATTGGTGACACCTGGAGGAGACCATAGACTGGTGGT 542
QY 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrCysLysProIle 200
DB 543 TACATGTTAGAGTGTGTGTCTTGGTAATGGAAAGAGGAGATGGACCTCGAAGCCCAT 602
QY 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
DB 603 GCTGAGAGAGTGTGTTGATCATGCTGGGACTTCTATGTGTGCGAGAAACGTTGGAG 662
QY 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
DB 663 AAGCCCTTACCAGGCTGGATGATGTAGATTGTACTTGGCTGGAGAAAGCAGCGGACGC 722
QY 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
DB 723 ATCATTTCGCACTTCTAGAAATAGATGCAAGATCAGACAGACAAAGACATCTATAGAATT 782
QY 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
DB 783 GGAGACACCTGGAGCAAGAGGATATTCAGAGAAACCTGCTCCAGTGCATCTGCACAGGC 842
QY 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
DB 843 AACGGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCAATCATCGAGCGGA 902
QY 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
DB 903 TCTGGCCCTTTCACCGATGTTGCTGCAGCTGTGTACCAACCGCAGCTTACCCCACT 962
QY 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
DB 963 CTTCCCTATGGCCACTGTGTCCAGACAGTGGTGTGTCTACTCTGTGGGATCAGTGG 1022
QY 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
DB 1023 CTGAAGACACAAAGAAATAAGCAAAATGCTTTGACGTCGCTGGGCAACGAGTCACTGC 1082
QY 361 GlnGluThrAlaValThrGlnThrTyrGlyLysSerAsnGlyGluProCysValLeu 380
DB 1083 CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTCTTA 1142
QY 381 ProPheThrTyrAsnAspArgThrAspSerThrThrSerAsnTyrGluGlnAspGlnLys 400
DB 1143 CCATTACCTACATGGCAGGACGTCGAGCACAACTTCAATATTAGCAGGACCAGAAA 1202
QY 401 TyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGly 420
DB 1203 TACTCTTCTGCACAGACCACTGTTTGTGTTTGTGCTCGAGGAGGAGAAATTCGAATGGT 1262
QY 421 AlaLeuCysHisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGlu 440
DB 1263 GCCTTGTGCCACTTCCCTTCCCTATACAACCAACCACTTACATGATGCTGCTGAG 1322
QY 441 GlyArgArgAspAsnMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLys 460
DB 1323 GGCAGAGAGACACATGAGTGGTGGGACCAACAGAACTATGATGCCGACAGAG 1382
QY 461 PheGlyPheCysProMetAlaAlaHisGluGluLeuCysThrThrAsnGluGlyValMet 480
DB 1383 TTTGGTCTTCCCTGCTGCCACGAGGAAATCTGCACAACCAATGAAGGGGTCTATG 1442
QY 481 TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThr 500
DB 1443 TACGCAATTGGAGATCAGTGGGTAAGCAGCATGACATGGGTTCACATGATGAGGTGCAGC 1502

Db 799 GAAGGGGTAGTCCTACAAGATTGGTGACACCTGGAGGAGACACATGAGACTGGTGGT 858
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrThrCysLysProIle 200
Db 859 TACATGTTAGAGTGTGTGTCTTGGTAATCGAAAGAGAGATGGACTCGCAAGCCCAT 918
Qy 201 AlaGluLysCysPheAspHisAlaGlyThrSerTyrValValGlyGluThrTyrGlu 220
Db 919 GCTGAGAAGTGTCTTGTATCATGCTGCTGGGACTTCTTATGTGGTGGGAAACGTGGAG 978
Qy 221 LysProTyrGlnGlyTyrMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 979 AAGCCCTACCAAGCTCGATGATGATGATTGTTACTTGCCTGGGAGAGGAGCGGACGC 1038
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 1039 ATCATCTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCTTAGAATT 1098
Qy 261 GlyAspThrTyrPheSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 1099 GGAGACACTGGAGCAAGAGATATTCAGAGAAACCTGCTCCAGTGGATCTGCACAGGC 1158
Qy 281 AsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1159 AACGCCGAGGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCCACATCAGCGGA 1218
Qy 301 SerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1219 TCTGGCCCTTACCGATGTTGCTGACGTGTTTACCAACCGCAGCCCTCACCCCGAGCCT 1278
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTyr 340
Db 1279 CTTCCCTATGGCACTGTGTACAGACAGTGTGTGTCTACTCTGTGGGATGCAGTGG 1338
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1339 CTGAAGACACAAGGAAATTAAGCAATGCTTTGCACGTGCTGGGCAACGGAGTCAGCTGC 1398
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1399 CAAGAGACAGCTGTAAACCGACTTACGTTGGCACTCAATGGAGAGCCATGTGTCTTA 1458
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1459 CCAITTCATCAATGGCAGGACGTTCTACTCTCTGCACCAAGAGGGCGACAGACCGA 1518
Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1519 CATCTTTGGTGCAGCACAACTTCGAATTATGAGCAGGACCAAGAAATACCTCTTCTGCACA 1578
Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
Db 1579 GACCACACTGTTTGGTTTCAGACTCGAGGAGGAAATTCCAATGGTGCCTTGTGCCACTTC 1638
Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn 445
Db 1639 CCCCCTCTATACACACACCAATTAACACTGATTCGACTTCTGAGGGCAGAGAGACAAAC 1698
Qy 446 MetLysTyrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1699 ATGAAGTGGTGGGACCAACACAGAACTATGATGCCGACCAAGATTTGGGTTCTGCCCC 1758
Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
Db 1759 ATGGCTGCCACAGAGAAATCTGCACAACCAATGAAGGGGTTCATGTCGCAATTCGAGAT 1818
Qy 486 GlnTyrAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db 1819 CAGTGGGATACAGCATGACATGGGTTCATGATGAGTGGTGCACGTGTGTGGGAATGGT 1878
Qy 506 ArgGlyGluThrThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1879 CGTGGGAAATGACATGCAATTCCTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATGAC 1938

Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
Db 1939 ATCATTTCAATGTGAACGACACATTTCCACAGCGTCATCAAGAGGGGCATGCTGAAC 1998
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTyrLysCysAspProValAspGlnCysGln 565
Db 1999 TGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAAGTGTATCCCGTCGACCAATGCCAG 2058
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTyrGluLysTyrValHisGly 585
Db 2059 GATTACAGAGACTGGAGCTTTTATCAAAATTGGAGATTTCATGGGAGAAAGTATGTCATGGT 2118
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTyrHisCysGlnProLeu 605
Db 2119 GTCAAGATACCAAGTGTACTGCTATGCCGTGGCATTTGGGAGTGGCATTTGCCAACCTTTA 2178
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2179 CAGACCTATCCAGCTCAAGTGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAG 2238
Qy 626 ProAsnSerHisProIleGlnTyrAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2239 CCCAACTCCCAACCCCATCCAGTGGATGCACCAACCCAGAACCTTGGATAC 2298
Qy 646 LeuArgTyrArgProValSerIleProProArgAsnLeuGlyTyr 660
Db 2299 CTCAGGTGGAGACTGTGATGATCCCAACCCAGAACCTTGGATAC 2343

RESULT 6

ADR66637
ID ADR66637 standard; DNA; 7242 BP.
XX AC ADR66637;
XX DT 02-DEC-2004 (first entry)
XX DE Human prostatic carcinoma derived DNA SEQ ID 149 #3.
XX KW human; cytostatic; diagnosis; prostatic cancer;
XX KW differential expression analysis; db.
XX OS Homo sapiens.
XX PN WO2004076614-A2.
XX PD 10-SEP-2004.
XX PF 22-FEB-2004; 2004WO-DE000433.
XX PR 27-FEB-2003; 2003DE-01009985.
XX PR 14-MAY-2003; 2003DE-01022134.
XX PA (HINZ/) HINZMANN B.
XX PA (DAHL/) DAHL E.
XX PA (ROSE/) ROSENTHAL A.
XX PA (HERM/) HERMANN K.
XX PA (PILA/) PILARSKY C.

XX Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarczyk C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
PI Xinzhang L, Staub E;
XX WPI; 2004-653386/63.
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX Claim 1; Page 1040-1042; 1607pp; German.

XX This invention describes novel cytostatic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer

Db 1773 CGTGGGAAATGGACATGATGCTTCCGACTTCGAGATCATGTCATTTGTGATGAC 1832
Qy 526 IleThrTyrAenValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAen 545
Db 1833 ATCACTTAAATGTGAACGACACATTCACAAGCGTCATGAAGAGGGGCACATGCTGAAC 1892
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrrpLysCysAspProValAspGlnCysGln 565
Db 1893 TGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAAAGTGTGATCCCGTCGACCAATGCCAG 1952
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrrpGluLysTyrValHisGly 585
Db 1953 GATTCAGAGATGGGACGCTTTTATCAAAATGGAGATTCATGGGAGAAGTATGTGCATGGT 2012
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrrpHisCysGlnProLeu 605
Db 2013 GTCAGATACCAAGTGTCTACTGTATGGCGTGGCAATGGGAGTGGCAATGCCAACCTTTA 2072
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2073 CAGACCTATCCAAGCTCAAGTGGTCTCTCGAAGTATTTATCACTGAGACTCCGAGTCAG 2132
Qy 626 ProAenSerHisProIleGlnTrrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2133 CCCAACTCCCAACCCCATCCAGTGGAAATGCACACAGCCCATCTCACATTTCCAAAGTACATT 2192
Qy 646 LeuArgTrrpArgPro 650
Db 2193 CTCAGGTGGAGACCT 2207

RESULT 7

AD65953
ID AD65953 standard; DNA; 7242 BP.
XX
AC AD65953;
XX
DT 02-DEC-2004 (first entry)
XX
XX Human prostatic carcinoma derived DNA SEQ ID 149 #1.
DE human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX
XX Homo sapiens.
OS
XX WO2004076614-A2.
XX
XX 10-SEP-2004.
XX
XX 22-FEB-2004; 2004WO-DE0000433.
XX
XX 27-FEB-2003; 2003DE-01009985.
PR
XX 14-MAY-2003; 2003DE-01022134.
XX
XX (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX
PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;
PI Schmitt A, Beckmann G, Bruemendorf T, Kimmennann H, Roepcke S;
PI Xinzhong L, Staub E;
XX
XX WPI; 2004-653386/63.
XX
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
XX agents.
XX
PS Claim 1; Page 385-386; 1607pp; German.
XX

CC This invention describes novel cytostatic polynucleotide and polypeptide
CC sequences which can be used in a method for diagnosing prostatic cancer
CC or the risk of developing prostatic cancer. Diagnosis is based on
CC determining over transcription or over expression of the sequences in
CC prostatic tissue. Screening for inhibitors of the sequences or detection
CC substances involves a binding assay, any compounds that bind are
CC selected, optionally after deconvolution of mixtures. Detection of a
CC predetermined minimum level of the reporter indicates the presence of a
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
CC short-interfering RNA or ribozymes; an organic molecule of molecular
CC weight below 5000, preferably 300, that binds to the polypeptide; an
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the
CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human
CC (monoclonal) antibody directed against Ab or any of the above derivatised
CC with a reporter group, cell toxin, immunostimulatory molecules and/or
CC radioisotope. The polynucleotides are identified in human prostatic
CC cancer by differential expression analysis, using DNA microarrays,
CC between normal and tumorous tissues, with (over)expression being detected
CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. AD65905-AD65954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.
XX

SQ Sequence 7242 BP; 1932 A; 1884 C; 1785 G; 1641 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.1e-239 Length: 7242
Score: 3693.50 Matches: 648
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 2
Query Match: 97.69% Indels: 15
DB: 13 Gaps: 1
US-09-581-651D-1 (1-660) x AD65953 (1-7242)
Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 213 AACTTGGTGGCACTTGCCTCCCGTGGGGCGTCTCTCCCCACCGCTCTCAACATGTT 272
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 273 AGGGTCCGGGGCCGGGCTGCTGCTGCCCGCTCCAGTGCCTGGGACAGCGGTGCC 332
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
Db 333 TCCACGGGAGGCTCGAAGAGCAGAGGCGAGCTCAGCAAAATGGTTTCAGCCCCAGTCCCG 392
Qy 61 ValAlaValSerGlnSerLysProGlyCysTrrpAspAsnGlyLysHisTyrGlnIleAen 80
Db 393 GTGGCTGTTCAGTCAAAGCAAGCCCGTGTATTGACAAATGGAAACACTATCAGTAAT 452
Qy 81 GlnGlnTrrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTrrpGlyGlySer 100
Db 453 CAACAGTGGGAGCGGACCTACTAGGCAATCGTGGTTGGTTTGTACTTGTATGAGGAGAAC 512
Qy 101 ArgGlyPheAenCysGluSerLysProGluAlaGluThrCysPheAspLysTrrpThr 120
Db 513 CGAGGTTTTAACTCGAGAGTAAACCTGAAGCTGAAGAGACTTGGCTTTGACAAGTACACT 572
Qy 121 GlyAenThrTrrpArgValGlyAspThrTrrpGluArgProLysAspSerMetIleTrrpAep 140
Db 573 GGGAAACATTACCGAGTGGTGACACTTATGAGGGTCTCTAAAGACTCCATGATCTGGGAC 632
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160

Db 633 TGTACTTCATCGGGCTGGCGAGGAGAAATAAGCTGTACCATCGCAACCGCTGCCAT 692
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTyrArgArgProHisGluThrGlyGly 180
Db 693 GAAGGGGTCTAGTCTCTACAGATTGGTGACACCTGGAGGAGACCATGAGACTGGTGGT 752
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrTyrCysLysProIle 200
Db 753 TACATGTTAGAGTGTGTCTTGGTAAATGGAAGAGGAGAAATGGACCTGCAAGCCATA 812
Qy 201 AlaGluLysCysPheAspHisAlaLeuGlyThrSerTyrValValGluThrTyrGlu 220
Db 813 GCTGAGAAAGTGTGGTATCATGCTGCTGGGACTTCTTATGTGGTGGGAGAACCTGGGAG 872
Qy 221 LysProTyrGlnGlyTyrMetMetValAspCysThrCysLeuGlyGluCysLysArg 240
Db 873 AAGCCCTACCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 932
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 933 ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACCAAGGACATCTATAGAATT 992
Qy 261 GlyAspThrTyrSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 993 GGAGACACTGGAGCAAGAGGATAATCGAGAAACCTGCTCCAGTGCATCTGCACACAGC 1052
Qy 281 AsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1053 AACGGCCGAGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGAGCGGA 1112
Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1113 TCTGGCCCTTACCGATGTTCTGTCAGCTGTTTACCAACCGCAGCTCCACCCCGCCT 1172
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
Db 1173 CCTCCCTATGGCCACTGTGTACAGACAGTGGTGTGGTCTACTCTGTGGGGATGCGTGG 1232
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1233 CTGAGACACAGGAAATAGCAAAATGCTTTGACAGTGCCTGGGCAACGGAGTCACTGC 1292
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
Db 1293 CAAGACACAGCTGTAAACCCAGACTTACGCTGGCAACTCAATGGAGAGCCATGTCTTA 1352
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1353 CCATTCACTCAATGGCAGGACGCTTCTACTCTGCACACAGAGGGCGACAGGACGGA 1412
Qy 389 ----- AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1413 CATCTTGGTGCAGCAACACTTCGAAATTATGAGCAGGACCAAGAAATCTCTTTCTGCACA 1472
Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
Db 1473 GACCACACTGTTTGGTTTCAGACTCGAGAGGAAATTCGAATGGTGGCTTGTGCCACTTC 1532
Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn 445
Db 1533 CCCTTCTCTATACAAACCAACCAATTTACACTCAITGTCACTTCTGAGGGGAGAGACAAAC 1592
Qy 446 MetLysTyrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1593 ATGAAGTGGTGGGACCAACACAGAACTATGATGCCGACCAAGAGTTTGGGTTCTGCCCC 1652
Qy 466 MetAlaHisGluGluLeuGlyCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
Db 1653 ATGGCTGCCACGAGGAAATCTGCACACCAATGAAGGGGTCAATGACCGCATTTGGAGAT 1712
Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db 1713 CAGTGGGATAGCAGCATGACATGGGTACATGATGATGAGGTGCACGTGTGTGGGAATGGT 1772

Qy 506 ArgGlyGluThrThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1773 CGTGGGGAATGGACATGCATTCCTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATCAC 1832
Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
Db 1833 ATCACTTACAAATGTGAACGACACATTCACAAAGCGTCAATGAAGAGGGGCACATGCTGAAC 1892
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg 565
Db 1893 TGTACATCTCTCGGTCAAGGTCGGGGCAGGTGGAGGTGTGATCCCGTCGACCAATGCCAG 1952
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTyrGluLysTyrValHisGly 585
Db 1953 GATTACAGAGACTGGGACGTTTTTATCAAAATGGAGATTCATGGGAGAGTATGTGTCATGCT 2012
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluThrHisCysGlyProLeu 605
Db 2013 GTCAGATACCACTGCTACTGCTATGGCGGTGGCAATGGGGAGTGGCAATGCCAACCTTTA 2072
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2073 CAGACCTATCCAACTCAAGTGGTCTGTGCGAGTATTTATCACTGAGACTCCGAGTCCG 2132
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2133 CCCAACTCCCAACCCCTCAGTGGATGCACACAGCCATCTCACATTTCCCAAGTACATT 2192
Qy 646 LeuArgTyrArgPro 650
Db 2193 CTCAGTGGAGACCT 2207
RESULT 8
ACC00412 ID ACC00412 standard; cDNA; 7550 BP.
XX AC ACC00412;
XX 04-JUL-2003 (first entry)
XX Human cell adhesion and extracellular matrix protein, CADECM-21, DNA.
DE Human; anti-HIV; anti-allergic; cerebroprotective; antiparkinsonian;
KW anticonvulsant; neurotropic; neuroprotective; immunosuppressive;
KW dermatological; anti-inflammatory; cytostatic; antiarteriosclerotic;
KW gene therapy; cell adhesion; extracellular matrix; CADECM;
KW immune system disorder; AIDS; allergy; neurological disorder; stroke;
KW Parkinson's disease; epilepsy; developmental disorder; Down's syndrome;
KW cerebral palsy; connective tissue disorder; systemic lupus erythematosus;
KW genetic disorder; Alport's syndrome; cell proliferative disorder; cancer;
KW atherosclerosis; gene; 88.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH CDS 271..7071
FT /*tag= a
FT /product= "CADECM-21"
XX
XX WO2003027230-A2.
XX
XX 03-APR-2003.
XX
XX 02-AUG-2002; 2002WO-US024649.
XX
XX 03-AUG-2001; 2001US-0309964P.
PR 03-AUG-2001; 2001US-0310119P.
PR 17-AUG-2001; 2001US-0313091P.
PR 31-AUG-2001; 2001US-0316771P.
PR 07-SEP-2001; 2001US-0317896P.
PR 21-SEP-2001; 2001US-0324781P.
PR 05-OCT-2001; 2001US-0327606P.

PR 12-OCT-2001; 2001US-0328960P.
PR 09-NOV-2001; 2001US-0344471P.
PR 17-MAY-2002; 2002US-0381291P.
XX (INCY-) INCYTE GENOMICS INC.
XX Burford N, Warren BA, Duggan BM, Mason PM, Richardson TW, Yue H;
PI Forsythe IJ, Elliott VS, Griffin JA, Gorvad AE, Azimzai Y;
PI Kallick DA, Xu Y, Honchell CD, Baughn MR, Gietzen KJ, Lee S;
PI Walia NK, Tang YT, Nguyen DB, Becha SD, Lee SY, Ramkumar J;
XX WPI; 2003-354645/33.
DR P-PSDB; ABR40124.
XX New human cell adhesion and extracellular matrix proteins (CADECM),
PT useful for diagnosing, treating or preventing disorders associated with
PT aberrant CADECM expression, e.g. cancer, AIDS, atherosclerosis, allergies
PT or stroke.
XX Claim 12; Page 231-233; 234pp; English.
XX The present invention relates to novel human cell adhesion and
CC extracellular matrix proteins, CADECM-1 to CADECM-22, and their coding
CC sequences (ACC00392-ACC00413 and ABR40104-ABR40125). The coding sequences
CC and proteins are useful in diagnosing, treating and preventing disorders
CC associated with aberrant expression of CADECM, such as immune system
CC disorders (e.g. AIDS or allergies), neurological disorders (e.g. stroke,
CC Parkinson's disease or epilepsy), developmental disorders (e.g. Down's
CC syndrome or cerebral palsy), connective tissue disorders (e.g. systemic
CC lupus erythematosus), genetic disorders (e.g. Alport's syndrome) or cell
CC proliferative disorders (e.g. cancer or atherosclerosis)
XX Sequence 7550 BP; 1996 A; 1993 C; 1899 G; 1672 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1.15e-239 Length: 7550
Score: 3693.50 Matches: 648
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 2
Query Match: 97.69% Indels: 15
DB: 8 Gaps: 1

US-09-581-651D-1 (1-660) x ACC00412 (1-7550)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 217 AACTTGGTGGACACTTGGCTCCCGTGGGGCTCTCTCCCCACCGCTCAACATGCTT 276

Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 277 AGGGTCCGGGCGCGGCTGCTGCTGGCGCTCCAGTGCCTGGGACAGCGGTGCC 336

Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
Db 337 TCCACGGAGCGCTCGAAGACGAAGAGGAGGCTCAGCAATGGTTTTCAGCCCGCCCG 396

Qy 61 ValAlaValSerGlnSerLysProGlyCysThrAspAsnGlyLysHisTyrGlnIleAsn 80
Db 397 GTGGCTGTCACTCAAGACGACCCCGTGTATGACATGGAAACACTATCAGATAAT 456

Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
Db 457 CAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGAGGAGC 516

Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluThrCysPheAspLysTyrThr 120
Db 517 CGAGGTGTTAACTCGCAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTCACAAAGTAC 576

Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 577 GGAACACTTACCGAGTGGGTGACACTTATGAGCGCTCTTAAGACTTCAATGATCTGGGAC 636

Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160

Db 637 TGTACCTGTCATCGGGGCTGGCGGAGGAGAATAAGCTGTACCATCGAAAAACCGCTGCAT 696
Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
Db 697 GAAGGGGGTCACTCTCAAGATTGGTGACACCTGGAGGAGACCAATAGACTGTGTGGT 756
Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysPheProle 200
Db 757 TACATGTTAGAGTGTGTGTCTTGGTAAATGGAAAGAGAGATGACCTGCAAGCCCAT 816
Qy 201 AlaGluLysCysPheAspHisAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db 817 GCTGAGAAGTGTGTTGATCATGCTGGGACCTTCTATGTGTGGAGAAACCTGGGAG 876
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 877 AGCCCTTACCAGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 936
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 937 ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTATAGAATT 996
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db 997 GGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGCACAGGC 1056
Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1057 AACGGCCGAGAGAGTGGAGTGTGAGAGGACACCTCTGTGACAGACACATCGAGCGGA 1116
Qy 301 SerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1117 TCTGGCCCTTCCCGATGTTGGTGAGCTGTTTACCAACCGGACGCTCACCCCGAGCT 1176
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
Db 1177 CCTCCCTATGGCCACTGTGTACACAGAGTGGTGGTGTACTCTGTGGGGATCGAGTGG 1236
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1237 CTGAAGACACAGGAAATAGCAATGCTTTGACAGTGCCTGGGCAACGGAGTCACTGC 1296
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyLysAsnSerAsnGlyGluProCysValLeu 380
Db 1297 CAAGACAGAGCTGTAACCCAGACTTACGGTGGCAACTCAATGGAGAGCCATGTGTCTTA 1356
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1357 CCATTCACCTACAAATGGCAGGACGCTTCTACTCTGCACCAAGGGCGGACAGACGGA 1416
Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1417 CATCTTTGGTGCACACAACTTCGAATATGAGCAGGACCAAGAAATACTCTTTCTGCACA 1476
Qy 406 AspHisThrValLeuValGlnThrArgGlyLysAsnSerAsnGlyAlaLeuCysHisPhe 425
Db 1477 GACCACACTGTTTGGTTTCAGACTCGAGGAGAAATCCAAATGCTGCTTGTGCCACTTC 1536
Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgAsnAspAsn 445
Db 1537 CCTTCTCTATACAAACCACTTACACTGATTTGCACTTCTCGAGGGGAGAGAGACAAAC 1596
Qy 446 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1597 ATGAAGTGGTGGGACCAACACAGAACTATGATGCCAGCCAGAGATTGGGTTCTGCCCC 1656
Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
Db 1657 ATGGCTGCCACGAGGAAATCTGCACAACTTGAAGGGGTGATGTATCCGCTTGGAGAT 1716
Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505

Db 1717 CAGTGGATAGGACGATGATGCGGTTCACATGATGAGGTGCAGCTGTGTTGGAAATGGT 1776
Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1777 CGTGGGAATGGACATGATGCTTCCCTACTCGCAGCTTCGAGATCAGTCATTTGTTGATGAC 1836
Qy 526 IleThrTyrAsnValAsnAspThrPheHisIlyArgHisGluGluGlyHisMetLeuAsn 545
Db 1837 ATCACTTACAAATGGAACGACACATTCACAAGCGCTCATGAAGAGGGGCGCATGCTGAAC 1896
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGly 565
Db 1897 TGTACATGCTTCGCTAGGGTGGGGGAGGTGGAGTGTGATCCCGTCGACCAATGCCAG 1956
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluTyrValHisGly 585
Db 1957 GATTACAGAGCTGGGACGTTTATCAAAATGGAGATTCTGGGAGAAATGATGTGATGGT 2016
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
Db 2017 GTCAGATACCAAGTCAAGTGTGCTGTGAAGTATTTATCACTGAGATCCGAGTCAG 2076
Qy 606 GlnThrTyrProSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2077 CAGACCTATCAAGCTCAAGTGTGCTGTGAAGTATTTATCACTGAGATCCGAGTCAG 2136
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2137 CCCAACTCCCAACCCATCCAGTGAATGCACCAAGCCATCTCACATTTCCAAAGTACATT 2196
Qy 646 LeuArgTrpArgPro 650
Db 2197 CTCAGTGGAGACCT 2211
RESULT 9
ADP64998
ID ADP64998 standard; DNA; 8027 BP.
XX
AC ADP64998;
DT
DT 12-AUG-2004 (first entry)
XX
XX Human fibronectin 1 (FNI), transcript variant 1 DNA sequence.
DE
XX
XX autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX
XX Homo sapiens.
XX
XX WO2003072827-A1.
XX
XX
XX 04-SEP-2003.
XX
XX 31-OCT-2002; 2002WO-US035433.
XX
XX 31-OCT-2001; 2001US-0336220P.
XX
XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
PA
XX
XX Hirsch R, Thorton SL;
XX
XX WPI; 2003-712740/67.
XX
XX GENBANK; NM_002026.
XX
XX Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX

PS Disclosure; Page; 56pp; English.
XX
XX The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthritides in the patient, where gene expression of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC induced arthritis. The compositions of the invention have the following
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This polynucleotide
CC represents a DNA sequence relating to the genes used in the analysis and
CC treatment of autoimmune diseases or arthritides. Note: This sequence is
CC not shown in the specification. It has been supplied in an electronic
CC format from WIPO.
XX
SQ Sequence 8027 BP; 2130 A; 2111 C; 1973 G; 1813 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,24e-239 Length: 8027
Score: 3693.50 Matches: 648
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 2
Query Match: 97.69% Indels: 15
DB: 11 Gaps: 1

US-09-581-651D-1 (1-660) x ADP64998 (1-8027)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
Db 214 AACTTGGTGGCAACTTGCCTCCCGGTGCGGGCTGCTCTCCCCACCGCTCAACATGCTT 273
Qy 21 ArgGlyProGlyProGlyLeuLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 274 AGGGGTCCGGGGCCCGGGTGTCTGCTGGCCGTCAGTGCCTGGGGACAGCGGTGCC 333
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
Db 334 TCCACGGGAGCCTCGAAGAGCAAGAGCGAGCTCAGCAATGGTTTCAGCCCGAGTCCCG 393
Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
Db 394 GTGGCTGTCACTCAAGCAAGCCCGGTGTTATGACAATGGAAACATCATAGATAAAT 453
Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
Db 454 CAACAGTGGGAGCGGACCTACCTAGGCATGCGTGGTTGTACTTGTATGGAGGAGC 513
Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120
Db 514 CGAGGTTTTAACTCGCAGAGAGTAACCTGAAGCTGAAGAGACTTGTCTTTGACAGTCACT 573
Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 574 GGAACACATTACCGAGTGGGTGACCTTATGACGCTCTTAAGAGCTCCATGATCTGGGAC 633
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
Db 634 TGTACCTGCAATCGGGCTGGCGGAGGAGATTAAGCTGTACCATTGCAAAACCGCTGCAT 693

QY 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
DB 694 GAAGGGGTGAGTCTCTACAGATTGGTGACACCTGGAGGAGACCATGAGACTGGTGGT 753
QY 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle 200
DB 754 TACATGTTAGAGTGTGTGTCTTGGTAATGGAAAAGGAGAATGACCTGCAAGCCCAT 813
QY 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
DB 814 GCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTATGTGGTGGGAGAACGTGGAG 873
QY 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
DB 874 AAGCCCTACCAAGGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 933
QY 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
DB 934 ATCACTTGCACTTCTAGAAAATAGATGCAACGATCAGGACACAAAGGACATCTATAGAATT 993
QY 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
DB 994 GGAGACACCTGGAGCAAGAAGGATAATCGAGGAAACCTGCTCCAGTGATCTGCACAGGC 1053
QY 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrSerSerGly 300
DB 1054 AAGGCCGAGGAGTGGAGTGTGAGAGGACACCTCTGTGCAGACACATCGAGCGGA 1113
QY 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
DB 1114 TCTGSCCCTTACCAGATGTTCTGTGACGTGTTTACCACCGCAGCCTACCCCGAGCCT 1173
QY 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
DB 1174 CTTCCCTATGSCCACTGTGTACAGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1233
QY 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
DB 1234 CTGAAGACACAGGAAATAGCAAAATGCTTTGCAGCTGCTGGGCAAGGAGTCAGCTGC 1293
QY 361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
DB 1294 CAAGAGACAGCTGTAAACAGACTTACGGTGGCAACTCAATGGAGAGCCATGTGTCTTA 1353
QY 381 ProPheThrTyrAsnAspArgThr----- 388
DB 1354 CCATTCACCTACAATGGCAGGACGTTCTACTCTCCACCGACCGAAGGCGCAGGACCGA 1413
QY 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
DB 1414 CATCTTTGGTCGACGACCAACTTCGAATATATGAGCAGGACCAAGAAATATCTTTCTGCACA 1473
QY 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
DB 1474 GACCACACTGTTTGTGTTTCAGACTCGAGAGGAAATTCCAATGGTGTGCTTGTGCCACTTC 1533
QY 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgArgAspAsn 445
DB 1534 CCCCCTCTATACAACAACCAAAATACACTGTGATTCGACTTCTGAGGGGAGAGAGACAAC 1593
QY 446 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheCysPro 465
DB 1594 ATGAAGTGTGTGGGACCAACAGAACTATGATGCGGACCAAGATTTGGGTTCGCCCC 1653
QY 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
DB 1654 ATGGCTGCCACAGGAAATCTGCACCAACCAATGAAGGGGTCACTGACCGCATTTGGAGAT 1713
QY 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
DB 1714 CAGTGGGATAAGCAGCATGACATGGGTACATGATGAGGTGACGTGTGTGGGAATGGT 1773
QY 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525

DB 1774 CTTGGGGAATGGACATGCATTGCTACTCTCCAGCTTCGAGATCAGTGCATTGTTGATGAC 1833
QY 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
DB 1834 ATCACTTACAAATGTGAACGACACATTCACAAGCGTCATGAAGAGGGGCACATGCTGAAC 1893
QY 546 CysThrCysPheGlyGlnGlyArgGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg 565
DB 1894 TGATACATGCTTCGGTCAAGGTTCGGGCGAGGTGGAGTGTGATCCCGTCGACCAATGCCAG 1953
QY 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly 585
DB 1954 GATTACAGAGACTGGAGCTTTTATCAAAATGGAGATTCATGGGAGAAAGTATGTGCATGGT 2013
QY 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
DB 2014 GTCAGATACCACTGCTACTGCTATGCCCGTGGCAATTCGGGAGTGGCATTCGCAACCTTTA 2073
QY 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
DB 2074 CAGACCTATCAAGCTCAAGTGTCTCTGCGAAGTATTTATCACTGAGACTCCGAGTCAG 2133
QY 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
DB 2134 CCCAACTCCCAACCCCATCCAGTGGATGCACACAGCCATCTCACATTTCCCAAGTACATT 2193
QY 646 LeuArgTrpArgPro 650
DB 2194 CTCAGGTGGAGACCT 2208
RESULT 10
ADG89565
ID ADG89565 standard; DNA; 8027 BP.
XX
AC ADG89565;
DT 11-MAR-2004 (first entry)
XX
DE Human fibronectin gene #3.
XX
KW metastatic cancer cell differentiation; mutated fibronectin;
KW metastatic cancer; human; ds; gene; fibronectin.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 268..7335
FT /*tag= a
FT /product= "Human fibronectin protein"
XX
PW WO2003100027-A2.
XX
PD 04-DEC-2003.
XX
PP 28-MAY-2003; 2003WO-US016736.
XX
PR 28-MAY-2002; 2002US-0383530P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
XX
PI Wang R;
XX
DR WPI; 2004-035134/03.
DR P-PSDB; ADG89560.
XX
PT Identifying a cell that differentiates into a metastatic cancer cell,
PT useful for preventing metastatic cancer, comprises identifying a mutated
PT fibronectin in the cell.
XX
PS Disclosure; SEQ ID NO 8; 137pp; English.
XX
CC The invention comprises a method for identifying a cell that will

CC differentiate into a metastatic cancer cell, the method involves
 CC identifying a mutated fibronectin in the cell. The method of the
 CC invention is useful for preventing metastatic cancer. The present DNA
 CC sequence represents a human fibronectin gene.

XX
 SQ Sequence 8027 BP; 2130 A; 2111 C; 1973 G; 1813 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.24e-239 Length: 8027
 Score: 3693.50 Matches: 648
 Percent Similarity: 97.44% Conservative: 0
 Best Local Similarity: 97.44% Mismatches: 2
 Query Match: 97.69% Indels: 15
 DB: 12 Gaps: 1

US-09-581-651D-1 (1-660) x ADG89565 (1-8027)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
 Db 214 AACTTGGTGGCAACTTGCCTCCCGTGGGGCGTCTCTCCCCACCGCTCTCAACATGCTT 273
 Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
 Db 274 AGGGGTCCGGGGCCGGGCTGCTGCTGCTGCGCGTCCAGTCCCTGGGACAGCGGTGCC 333
 Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60
 Db 334 TCCACGGGAGCCTCGAAGAGCAGAGGCGAGCTCAGCAAAATGGTTCCGCCCCAGTCCCG 393
 Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrrAspAsnGlyLysHisTyrrGlnIleAsn 80
 Db 394 GTGGCTGTGACGTAAGCAAGCCGGTGTGTATGACAAATGGAAACACATATCAGATAAT 453
 Qy 81 GlnGlnTrpGluArgThrTyrrLeuGlyAsnAlaLeuValCysTyrrGlyGlySer 100
 Db 454 CAACAGTGGGAGCGGACCTACTAGGCAATCGCTGGTGTGTATGAGAGGAGC 513
 Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGlnThrCysPheAspLysThr 120
 Db 514 CGAGGTTTTAATCGCAGAGTAAACCTGAAAGCTGAAAGAGACTTCTTGTGACAACTACT 573
 Qy 121 GlyAsnThrTyrrArgValGlyAspThrTyrrGluArgProLysAspSerMetIleTrpAsp 140
 Db 574 GGGAAACCTTACCAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGGAC 633
 Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
 Db 634 TGTACCTGCATCGGGGCTGGGCGAGGAGAAATAAGCTGTACATCGCAAAACCGCTGCCAT 693
 Qy 161 GluGlyGlnSerTyrrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
 Db 694 GAAGGGGTGAGTCTCAACAGATTGGTGACACCTGGAGGAGACACATGAGACTGGTGT 753
 Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle 200
 Db 754 TACATGTTAGAGTGTGTGTCTTGGTAAATGGAAAAGGAGAATGGAGCTGCAAGCCCAT 813
 Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrrValValGlyGluThrTrpGlu 220
 Db 814 GCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTTATGTGGTGGGAAACGTGGAG 873
 Qy 221 LysProTyrrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
 Db 874 AAGCCCTACCAAGGCTGGATGATGATGTTGACTTGCCTGGGAGAGCGGCGGACGC 933
 Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrrArgIle 260
 Db 934 ATCACTTGCACTTCTAGAAAATAGATGCAACCGATCAGGACACACAGGACATCTATAGAAT 993
 Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
 Db 994 GGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGATCTGCACAGGC 1053

Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
 Db 1054 AACGGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACACACATCGAGCGGA 1113
 Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrrGlnProGlnProHisProGlnPro 320
 Db 1114 TCTGGCCCTTACCAGATGTTGTCAGAGCTGTTTACCACCGCAGCTTACCCCGAGCT 1173
 Qy 321 ProProTyrrGlyHisCysValThrAspSerGlyValValTyrrSerValGlyMetGlnTrp 340
 Db 1174 CCTCCCTATGGCCACTGTGTACACAGACAGTGGTGTGTCTACTCTGTGGGATGCAGTGG 1233
 Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
 Db 1234 CTGAAGACACAAAGGAAATAAGCAAAATGCTTTGGACGTGCTCTGGGCAACGGAGTCAGTGC 1293
 Qy 361 GlnGluThrAlaValThrGlnThrTyrrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
 Db 1294 CAAGAGACAGCTGTAACCCAGACTTACGGTGGCACTCAATATGGAGAGCCATGTCTT 1353
 Qy 381 ProPheThrTyrrAsnAspArgThr----- 388
 Db 1354 CCATTCACTACAAATGCGAGGAGCTTCTACTCTGCACACCGAAGGCGGACAGCGGA 1413
 Qy 389 -----AspSerThrThrSerAsnTyrrGluGlnAspGlnLysTyrrSerPheCysThr 405
 Db 1414 CATCTTTGGTGCACACAACTTCGAATATATGAGCAGGACAGAAATATCTTTCTGCACA 1473
 Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
 Db 1474 GACCACACTGTTTGGTTCAGACTCGAGGAGGAAATTCGAATGTGTGCTTGTGCCACTTC 1533
 Qy 426 ProPheLeuTyrrAsnAsnHisAsnTyrrThrAspCysThrSerGluGlyArgArgAspAsn 445
 Db 1534 CCCTTCTTATACAAACACCACTTACACTGATTTGCATCTCTGAGGCGCAGAGAGACAAC 1593
 Qy 446 MetLysTrpCysGlyThrThrGlnAsnTyrrAspAlaAspGlnLysPheGlyPheCysPro 465
 Db 1594 ATGAAGTGGTGTGGGACACACAGAACTATGATGCCGACCAAGAGTTTGGGTTCTGCCCC 1653
 Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrrArgIleGlyAsp 485
 Db 1654 ATGGCTCCCAAGGAGAAATCTGCACAAACCAATGAAGGGGTGATGTACCGCATTTGGAGAT 1713
 Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
 Db 1714 CAGTGGATTAAGCAGCATGACATGGGTACATGATGAGTGCAGCTGTGTGGGAATGTT 1773
 Qy 506 ArgGlyGluTrpThrCysIleAlaTyrrSerGlnLeuArgAspGlnCysIleValAspAsp 525
 Db 1774 CGTGGGAAATGGACATGCATTTGCTACTCGCAGCTTCGAGATCAGTGCATTTGTGATGAC 1833
 Qy 526 IleThrTyrrAsnValAsnAspThrPheHisArgHisGluGluGlyHisMetLeuAsn 545
 Db 1834 ATCACTTACAAATGTAACGACACATTCACAAGCGTCATGAAGAGGGGACATCTCTGAAC 1893
 Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
 Db 1894 TGTACATGCTTCGGTCAAGGTCGGGCGAGGTGGAGTGTGATCCCGTCGACCAATGCCAG 1953
 Qy 566 AspSerGluThrGlyThrPheTyrrGlnIleGlyAspSerTrpGluLysTyrrValHisGly 585
 Db 1954 GATTCAGAGACTGGACGTTTTTATCAAAATGGAGATTCATGGGAGAGATGATGTGATGCT 2013
 Qy 586 ValArgTyrrGlnCysTyrrCysTyrrGlyArgGlyIleGlyGluTrpHisCysGlyProLeu 605
 Db 2014 GTCAATACCAAGTGTACTGCTATGGCCGTGGCATTTGGGAGTGGCATTCGCAACCTT 2073
 Qy 606 GlnThrTyrrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
 Db 2074 CAGACCTATCAAGCTCAAGTGTCTCTGTGAGATATTTATCTACTGAGACTCCGAGTCAG 2133
 Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrrIle 645

Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
 Db 1414 CATCTTTGGTGCAGCACAACTTCGAATTTATGAGCAGGACCAAGAAATCTCTTTCTGCACA 1473
 Qy 406 AspHisThrValLeuValGlnThrArgGlyCysSerAsnGlyAlaLeuCysHisPhe 425
 Db 1474 GACCACACTGTTTGGTTTCAGACTCGAGGAGGAATTCGAATGTTGGCTTGTGGCACTTC 1533
 Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgAspAsn 445
 Db 1534 CCCTTCCTATACACACCAATTCACACTGATTCGACTTCTGAGGGCAGAGAGACAC 1593
 Qy 446 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
 Db 1594 ATGAAGTGGTGTGGGACCAACAGAACTATGATGCCAGCACGAAGATTGGGTTCTGCCCC 1653
 Qy 466 MetAlaAlaHisGluGluLeuValThrAsnGluGlyValMetTyrArgIleGlyAsp 485
 Db 1654 ATGGCTGCCACAGGAAATCTGCACCAACCAATGAAGGGTTCATGTACCGCATTTGGAGAT 1713
 Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
 Db 1714 CAGTGGGATAGCAGCATGACATGGTGCATGATGAGTGCAGTGTGTGGGAATGGT 1773
 Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
 Db 1774 CGTGGGAATGGACATGCTGCTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATGAC 1833
 Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
 Db 1834 ATCACTTACATGTGAACGACACATTCACAAGCGTTCATGAAGAGGGGCACATGCTGAAC 1893
 Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
 Db 1894 TGTACATGCTTCGGTTCAGGTCGGGGCAGGTGGAGTGTGATCCCGTCGACCAATGCCAG 1953
 Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly 585
 Db 1954 GATTACAGAGCTGGGACGTTTTATCAAAATTCGAGATTCATGGGAGAAATGTCATGGT 2013
 Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
 Db 2014 GTCAATACCACTGCTACTGCTATGGCGTGGCATTTGGGAGTGGCATTTGCCAACCTTTA 2073
 Qy 606 GlnThrTyrProSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
 Db 2074 CAGACCTATCCAGCTCAAGTGGTCTGTGGAAGTATTTATCACTGAGACTCCGAGTCAG 2133
 Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
 Db 2134 CCCAACTCCCAACCCCATCCAGTGAATGCACACAGCCATCTCACATTTCCAGTACATT 2193
 Qy 646 LeuArgTrpArgPro 650
 Db 2194 CTCAGGTGGAGACCT 2208
 RESULT 12
 ADR67200
 ID ADR67200 standard; DNA; 8027 BP.
 XX
 AC ADR67200;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Human bladder cancer associated nucleotide sequence.
 XX
 KW bladder cancer tissue; bladder cancer; cytostatic; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN W02004076613-A2.
 XX

PD 10-SEP-2004.

XX 24-FEB-2004; 2004WO-DE000364.

XX 26-FEB-2003; 2003DB-01009729.

XX (HERR/) HERR A.

PA (HINZ/) HINZMANN B.

PA (DAHL/) DAHL E.

PA (STAU/) STAUB E.

PA (PILA/) PILARSKY C.

PA (SPEC/) SPECHT T.

XX Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;

XX WPI; 2004-653385/63.

XX New nucleic acids, and encoded proteins, from bladder cancer tissue,

XX useful for diagnosis, treatment and in screening for specific binding

XX agents.

XX Claim 1; Fig 3; 112pp; German.

XX The present invention describes nucleic acids (I) associated with bladder

XX cancer tissue. Also described: (1) peptides and proteins (II) containing

XX an amino acid sequence encoded by (I); (2) a method for diagnosing

XX bladder cancer (BC), or monitoring its progression, that uses (I), (II)

XX or agents (2) that inhibit, or bind to, (I) or (II), labelled with a

XX reporter; and (3) a method for treating BC that uses (I), (II) or (2).

XX (I) and (II) have cytostatic activity. (I) and (II) can be used to detect

XX (and monitor progression of) bladder cancer (BC), or the risk of

XX developing it; to screen for specific binding agents (2), and to treat

XX BC. (2) are also useful as diagnostic and therapeutic agents. The present

XX sequence represents a human nucleotide sequence associated with bladder

XX cancer, which is used in the exemplification of the present invention.

XX SQ Sequence 8027 BP; 2130 A; 2111 C; 1973 G; 1813 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1-24e-239 Length: 8027

Score: 3693.50 Matches: 648

Percent Similarity: 97.44% Conservative: 0

Best Local Similarity: 97.44% Mismatches: 2

Query Match: 97.69% Indels: 15

DB: 13 Gaps: 1

US-09-581-651D-1 (1-660) x ADR67200 (1-8027)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20

Db 214 AACTTGGTGGCAACTTGGCTCCCGGTGGGGCGTCTCTCCCCACCGCTCAACATGCTT 273

Qy 21 ArgGlyProGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40

Db 274 AGGGGTCCGGGGCCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 333

Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGlnSerPro 60

Db 334 TCACGGGAGCCCTCGAAGACCAAGAGCAGCGCTCAGCAATGTTTCAGCCGCCGCTCCCG 393

Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80

Db 394 GTGGCTGTCAAGTCAAGAGCCCGGTTGTTATGACAATGGAACACATATCAGATNAAT 453

Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100

Db 454 CAACAGTGGGAGCGGACCTTACCTAGGCAATGCGTGGTGGTGGTGGTGGTGGTGGTGG 513

Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120

Db 514 CGAGGTTTAACTCCGAGAGTAACTGAAGAGTGAAGAGACTTGTCTTTGACAGTACACT 573

Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140

|||||
Db 574 GGGAAACCTTACCGAGTGGGTGACACCTATGAGCGCTCTTAAAGACTCCATGATCTGGGAC 633
Qy 141 CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAenArgCysHis 160
Db 634 TGTACCTGCATCGGGCTGGCGAGGAGGAATAAGCTGTACCATCGCAAAACCGCTGCCAT 693
Qy 161 GluGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180
Db 694 GAAGGGGGTCTAGTCTCTAAGATTGGTGACACCTCGAGGAGACACATGAGACTGGTGGT 753
Qy 181 TyrMetLeuGluCysValCysLeuGlyAenGlyLysGlyGluTrpThrCysLysProIle 200
Db 754 TACATGTTAGAGTGTGTGTCTTGGTAATGGAAAAGAGNAATGGACCTCGAAGCCCAT 813
Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu 220
Db 814 GCTGAGAAGTGTCTTGTATCATGCTGCTGGGACTTCTTATGTGGTGGGAGAACCTGGGAG 873
Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
Db 874 AAGCCCTACCAAGCTCGATGATGGTAGATTGTACTTGCCTGGGAGAGGCGGACGC 933
Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
Db 934 ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAGACATCTCTATAGAAAT 993
Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAenLeuLeuGlnCysIleCysThrGly 280
Db 994 GGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1053
Qy 281 AenGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
Db 1054 AACGGCCGAGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAACACATCAGAGCGGA 1113
Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
Db 1114 TCTGGCCCTTCAACGATGTTCTGTGCAGCTGTTTACCAACCCGACGCTCACCCCCAGCCT 1173
Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
Db 1174 CCTCCCTATGCCACTGTGTGCACAGACAGTGGTGTGTCTACTCTGTGGGATGCACTGG 1233
Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db 1234 CTGAAGACACAAAGGAATAAAGCAATGCTTTGACAGCTGCTGGGCAACGGAGTCAGCTGC 1293
Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAenSerAenGlyGluProCysValLeu 380
Db 1294 CAAGGACAGCTGTMACCAGACTTACCGTGGCAACTCAATGGAGAGCCATGTCTTA 1353
Qy 381 ProPheThrTyrAsnAspArgThr----- 388
Db 1354 CCATTCCACTCAATGGCAGGACGTTCTACTCTGCACACAGGAAGGCGCAGGACGGA 1413
Qy 389 ----- AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
Db 1414 CATCTTTGGTGCAGCACCACTTCGAATATATGAGCAGGACCAAAATATCTTTCTGCACA 1473
Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAenSerAenGlyAlaLeuCysHisPhe 425
Db 1474 GACCAACTGTTTGGTTCAGACTCGAGGAGGAATTCGAATGGTGCCTTGTGCCACTTC 1533
Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerSerGluGlyArgArgAspAen 445
Db 1534 CCTTCTCTATACAACAACCAATATACACTGATTCGACTTCTGAGGGCAGAGAGACAAAC 1593
Qy 446 MetLysThrCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
Db 1594 ATGAAGTGGTGGGACCAACAGAACTATGATGCCGACCAAGATTTGGGTTCTGCCCC 1653
Qy 466 MetAlaAlaHisGluGluCysThrThrAenGluGlyValMetTyrArgIleGlyAsp 485
|||||

1654 ATGGCTGCCACCAGGAAATCTGCACACCAATGAAGGGGTCAATGATCCGCATTGGAGAT 1713
Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAenGly 505
Db 1714 CAGTGGGATAAGCAGCATGACATGGGTCAATGATGAGGTGCACGTGTGTGGGAATGGT 1773
Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1774 CGTGGGGATGGACATGCATTCCTACTCGACGTTTCGAGATCAGTGCATTTGTTGATGAC 1833
Qy 526 IleThrTyrAsnValAenAspThrPheHisLysArgHisGluGlyHisMetLeuAen 545
Db 1834 ATCATTCAATCTGAACGACACATTCACACAGCGTCATGAAGAGGGGCACTGCTGAAC 1893
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
Db 1894 TGTACATGCTTCGGTCCGGGTCGGGCGAGGTGAAGTGTGATCCCGTCGACCAATGCCAG 1953
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly 585
Db 1954 GATTGAGAGACTGGACGTTTTATCAATTTGGAGATTCATGGGAGAAGTATGTGATGCT 2013
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
Db 2014 GTCAGATACCACTGCTACTGCTATGGCGTGGCATTTGGGAGTGGCATTCGCAACCTTTA 2073
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2074 CAGACCTATCCAAAGCTCAAGTGTCTGTGCAAGTATTTATCACTGAGACTCCGAGTCAG 2133
Qy 626 ProAenSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2134 CCNACTCCCAACCCCATCCAGTGGATGCACACAGCCATCTCACATTTCCAAATGATATT 2193
Qy 646 LeuArgTrpArgPro 650
Db 2194 CTCAGTGGAGACCT 2208
RESULT 13
ADQ38578
ID ADQ38578 standard; DNA; 6510 BP.
XX
XX AC ADQ38578;
XX
XX 18-NOV-2004 (first entry)
XX
XX DE Human SNP containing myocardial infarction-associated gene, SEQ ID 241.
XX
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX
XX KW cardiant; gene therapy; human; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2004058052-A2.
XX
XX PD 15-JUL-2004.
XX
XX PF 22-DEC-2003; 2003WO-US040978.
XX
XX PR 20-DEC-2002; 2002US-0434778P.
XX
XX PR 10-MAR-2003; 2003US-0453135P.
XX
XX PR 30-APR-2003; 2003US-0466412P.
XX
XX PR 23-SEP-2003; 2003US-0504955P.
XX
XX PA (APPL-) APPLERA CORP.
XX
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX
XX DR WPI; 2004-533949/51.
XX
XX DR P-PSDB; ADQ39406.
XX
XX PT Identifying an individual who has an altered risk for developing
myocardial infarction by detecting a single nucleotide polymorphism in

Db 1819 CAGTGGGNTAAGCAGCATGATGGGTCCATGATGAGGTCCACGTGTGTGGGAATGGT 1878
Qy 506 ArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
Db 1879 CGTGGGAATGGACATGATGCTTACTCGCAGCTTCGAGATCAGTGCATTTGTTGATGAC 1938
Qy 526 IleThrTyrAonValAenAspThrPheHisIysArgHisGluGluGlyHisMetLeuAen 545
Db 1939 ATCACTTACATGATGGAACGACACATTCACAAAGCGTCATGAAGAGGGGCACATGCTGAAC 1998
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
Db 1999 TGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAGTGTGATCCCGTCGACCAATGCCAG 2058
Qy 566 AspSerGluThrGlyThrPheTyrClnIleGlyAspSerTrpGluLysTyrValHisGly 585
Db 2059 GATTTCAGAGACTGGGACGCTTTTATCMAATTGGAGATTTCATGGGAGAAATGTCATGGT 2118
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
Db 2119 GTCAGATACCACTGCTACTGTATGCGCGTGGCATTTGGGGAGTGGCATTCGCAACCTTTA 2178
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2179 CAGACCTATCCAGCTCAAGTGGTCTCTCGAGTATTTATCACTGAGACTCCGAGTCAG 2238
Qy 626 ProAenSerHisProIleGlnTrpAenAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2239 CCCAACTCCCAACCCCATCCAGTGGAAATGCACACAGCCATCTCACATTTCCAAAGTACATT 2298
Qy 646 LeuArgTrpArgPro 650
Db 2299 CTCAGGTGGAGACCT 2313

RESULT 14
ADQ38582
ID ADQ38582 standard; DNA; 7823 BP.
XX
XX AC ADQ38582;
XX AC
XX DT 18-NOV-2004 (first entry)
XX DE Human SNP containing myocardial infarction-associated gene, SEQ ID 245.
XX DE
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human; gene; ds.
XX OS Homo sapiens.
XX FN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
XX DR P-PSDB; ADQ39410.
XX DR
XX PT Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.
XX PS Claim 7; SEQ ID NO 245; 145pp; English.
XX PS

CC The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting a SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.

XX
SQ Sequence 7823 BP; 2066 A; 2031 C; 1935 G; 1776 T; 0 U; 15 Other;

Alignment Scores:

Pred. No.: 3,06e-239 Length: 7823
Score: 3687.50 Matches: 647
Percent Similarity: 97.29% Conservatives: 0
Best Local Similarity: 97.29% Mismatches: 3
Query Match: 97.53% Indels: 15
DB: Gaps: 1

US-09-581-651D-1 (1-660) x ADQ38582 (1-7823)

Qy 1 AenLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAenMetIeu 20
Db 319 AACCTGGTGGCAACTTGCCTCCCGTCCGGTCCGGCGCTCTCCCCACCGCTCTCAACATGCTT 378
Qy 21 ArgGlyProGlyProGlyLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
Db 379 AGGGGTCCGGGCCCCGGCTGCTGCTGCGCCCTCCWGTGCTTGGGACAGCGGTGCC 438
Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
Db 439 TCCACGGGAGCCTCGAAGAGCAGAGGAGGAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGC 498
Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAenGlyLysHisTyrGlnIleAen 80
Db 499 GTGGCTGTCACTCAAGCAAGCCCGTGTGTATGACAAATGGAAACACTATCAGATAAAT 558
Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAenAlaLeuValCysThrCysTyrGlyGlySer 100
Db 559 CAACAGTGGGAGCGGACCTACCTAGGCAATCGGTGGTTTGTACTTTGTTATGGAGGAAGC 618
Qy 101 ArgGlyPheAenCysGluSerLysProGluAlaGluThrCysPheAspLysTyrThr 120
Db 619 CGAGGTTTAACTCGGAGAGTAAACCTGAAGAGACTTCTGTTTGACAAGTACACT 678
Qy 121 GlyAenThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
Db 679 GGGGAACACTTACCGAGTGGGTGACACTTATGAGCGTCTCTAAAGACTCCATCATCTGGGAC 738
Qy 141 CysThrCysIleGlyAlaGlyArgGlySerCysThrIleAlaAenArgCysHis 160
Db 739 TGTACCTGCATYGGGGCTGGGCGGAGGAGAAATAGCTGTACCATCGCAACCGTCCCAT 798
Qy 161 GluGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly 180

Db 799 GAAGGGGTAGTCTCTCAAGATTGGTGACACCTGGAGGAGACACATGAGACTGGTGGT 858
Qy TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluThrPThrCysLysProIle 200
Db TACATGTTAGAGTGTGTGTCTTGGTAATGGAAGAGGAGATGGACTCGAAGCCCAT 918
Qy AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyValValGlyGluThrTrpGlu 220
Db GCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTATGTGGTCCGAGAAACGTGGAG 978
Qy LysProTyGlnGlyTrpMetMetValAspCysVserCysLeuGlyGluGlySerGlyArg 240
Db AAGCCCTACCAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
Qy IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyArgIle 260
Db ATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACACAGACATCTCTATAGAA 1098
Qy GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
Db GGAGACACCTGGAGCAAGAGATAATCGAGAAACCTGCTCCAGTGCATCTGCACAGGC 1158
Qy AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrSerSerGly 300
Db AACGCCGAGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACACCATCAGAGCGGA 1218
Qy SerGlyProPheThrAspValArgAlaValTyGlnProGlnProHisProGlnPro 320
Db TCTGCCCCCTTACCGATGTTCTGTGCAGCTGTTTACCAACCGCAGCCTCACCCACAGCT 1278
Qy ProProTyGlyHisCysValThrAspSerGlyValValTySerValGlyMetGlnTrp 340
Db CCTCCCTATGCCACTGTGTACAGACAGTGGTGTGTGTCTACTCTGTGGGATGCACTGG 1338
Qy LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
Db CTGAAGACACAAAGGAAATAAGCAAATGCTTTGACAGTGCCTGGGCAACGGAGTCAGCTGC 1398
Qy GlnGluThrAlaValThrGlnThrTyGlyLysAsnSerAsnGlyGluProCysValLeu 380
Db CAAAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAATGGAGAGCCATGTGTCTTA 1458
Qy ProPheThrTyAsnAspArgThr----- 388
Db CCATTCCACTACATGGCAGGAGCTTCTACTCTCTGCACACAGAAAGGCGAGCAGCGGA 1518
Qy -----AspSerThrThrSerAsnTyGluGlnAspGlnLysTySerPheCysThr 405
Db CATCTTTGGTGCAGCAACAACCTTCGAATTTATGAGCAGGACACAGAAATACTCTTTCTGCACA 1578
Qy AspHisThrValLeuValGlnThrArgGlyLysAsnSerAsnGlyAlaLeuCysHisPhe 425
Db GACCACACTGTTTGGTTCAGACTCGAGAGGAAATTCGAATGTTGCTTGTGCCACTTC 1638
Qy ProPheLeuTyAsnAsnHisAsnTyThrAspCysThrSerGluGlyArgArgAspAsn 445
Db CCCTTCTCTATACACACCAACCAATACACTGATTCGACTTCTGAGGGCAGAGAGACAC 1698
Qy MetLysTrpCysGlyThrThrGlnAsnTyAspAlaAspGlnLysPheGlyPheCysPro 465
Db ATGAAGTGGTGTGGGACACACAGAACTATGATGCCAGCAGAAAGTTTGGGTTCTGCCCC 1758
Qy MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyArgIleGlyAsp 485
Db ATGGCTGCCACAGGAGAAATCTGCACACCAACCAATGAAGGGGTCATGTCGCGATTGGAGAT 1818
Qy GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
Db CAGTGGGNTAAGCAGCATGACATGGGTACATGATGAGTGCAGCTGTGTGGGAATGGT 1878
Qy ArgGlyGluTrpThrCysIleAlaTySerGlnLeuArgAspGlnCysIleValAspAsp 525

Db 1879 CGTGGGGAATGGACATGCATTCCTACTCGAGCTTCGAGATCAGTCATGTTGATGAC 1938
Qy 526 IleThrTyAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545
Db 1939 ATCACTTACAAATGTGAACGACACATTCACAAGCGTCATGAAGAGGGGCACATGCTGAAC 1998
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspProValAspGlnCysGln 565
Db 1999 TGTACATGCTTCGGTTCAGGTCGGGCGAGGTGGAAAGTGTATCCCGTCGACCAATGCCAG 2058
Qy 566 AspSerGluThrGlyThrPheTyGlnIleGlyAspSerTrpGluLysTyValHisGly 585
Db 2059 GATTCAAGACTGGGAGCTTTTATCAANTTGGAGATTCATGGGAGAGTATGTGCATGGT 2118
Qy 586 ValArgTyGlnCysTyCysTyArgGlyArgGlyGlyGluTrpHisCysGlnProLeu 605
Db 2119 GTCAGATACCACTGCTACTGCTATGCGCGTGGCATTTGGGAGTGGCATTTGCCAACCTTTA 2178
Qy 606 GlnThrTyProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2179 CAGACCTTATCAAGCTCAAGTGGTCTCTGTCGAAGTATTTATCACTGAGACTCCGAGTCAG 2238
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyIle 645
Db 2239 CCCAACTCCCAACCCATCCAGTGGATGCCACAGCCATCTCACATTTCCAAATACATT 2298
Qy 646 LeuArgTrpArgPro 650
Db 2299 CTCAGGTGGAGACT 2313
RESULT 15
ADQ38583
ID ADQ38583 standard; DNA; 7848 BP.
XX AC ADQ38583;
XX DT 18-NOV-2004 (first entry)
XX DE Human SNP containing myocardial infarction-associated gene, SEQ ID 246.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human; gene; db.
XX OS Homo sapiens.
XX OS WO2004058052-A2.
XX PN 15-JUL-2004.
XX PD 22-DEC-2003; 2003WO-US040978.
XX PF 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Takoubova O;
XX WIPI; 2004-533949/51.
XX DR P-PSDB; ADQ39411.
XX PT Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.
XX PS Claim 7; SEQ ID NO 246; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
XX CC has an altered risk for developing myocardial infarction. The method
XX CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
XX CC the nucleotide sequences given in the specification in the individual's

CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC polynucleotide sequence represents a human myocardial infarction-
 CC associated gene containing one or more SNP's of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.
 XX
 SQ Sequence 7848 BP; 2069 A; 2046 C; 1939 G; 1778 T; 0 U; 16 Other;

Alignment Scores:
 Pred. No.: 3,07e-239 Length: 7848
 Score: 3687.50 Matches: 647
 Percent Similarity: 97.29% Conservative: 0
 Best Local Similarity: 97.29% Mismatches: 3
 Query Match: 97.53% Indels: 15
 DB: 13 Gaps: 1

US-09-581-651D-1 (1-660) x ADQ38583 (1-7848)

Qy 1 AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu 20
 Db 319 AACTTGGTGGCACTTGGCTCCCGTGGGGGTCTCTCCCGCCGCTCTCAATGCTT 378
 Qy 21 ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro 40
 Db 379 AGGGTCCGGGGCCGGGCTGCTGCTGGCCGTCWGTGCTGGGACAGCGGTGCC 438
 Qy 41 SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro 60
 Db 439 TCCACGGAGGCTCGAAGAGCAAGAGGAGGAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGC 498
 Qy 61 ValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGlnIleAsn 80
 Db 499 GTGGCTGTCAAGCAAGCAAGCCGGTGTGTATGCAATGGNAACACTATCAGATAAT 558
 Qy 81 GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGlyGlySer 100
 Db 559 CAACAGTGGAGCGGACCTACTAGGCAATGCGTTGGTTGTACTTTGATGGAGGAAGC 618
 Qy 101 ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysTyrThr 120
 Db 619 CGAGGTTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAGTACACT 678
 Qy 121 GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp 140
 Db 679 GGGAAACACTTACCGAGTGGGTGACACTTATGAGCGTCTCTAAAGACTCCCATGATCTGGGAC 738
 Qy 141 CysThrCysIleGlyValaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis 160
 Db 739 TGTACTCTGCATYGGGGCTGGGCGAGGAGATAAAGCTGTACCATCGCAAAACCGCTGCCAT 798
 Qy 161 GluGlyGlyGlnSerTyrLysIleGlyAspThrTyrArgArgProHisGluThrGlyGly 180
 Db 799 GAAGGGGTGCTGCTCTACAAGATTGGTGACACCTCGAGGAGACCATGAGACTGGTGT 858
 Qy 181 TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTyrThrCysLysProIle 200

Db 859 TACATGTTAGAGTGTGTGTCTTGTGTAATGGAAGAGAGATGGACCTCTCAAGCCCAT 918
 Qy 201 AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTyrGlu 220
 Db 919 GCTGAGAAGTGTGTTGATCATGTCTGGGACTTCTTATGTGTGGAGAAACAGTGGAG 978
 Qy 221 LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg 240
 Db 979 AAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAAGCAGCGGACGC 1038
 Qy 241 IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle 260
 Db 1039 ATCATTGTCATCTTCAAAATAGATGCAACGATCAGGACACAAGGACATCTCTATAGATT 1098
 Qy 261 GlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly 280
 Db 1099 GGAAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC 1158
 Qy 281 AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly 300
 Db 1159 AAGCGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTCTGTCAGACACCATCGAGCGGA 1218
 Qy 301 SerGlyProPheThrAspValArgAlaAlaValTyrGlnProGlnProHisProGlnPro 320
 Db 1219 TCTGGCCCTTCCCGATGTTCTGTCAGCTGTTTACCACCGCAGCCTCACCACGCT 1278
 Qy 321 ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp 340
 Db 1279 CCTCCCTATGGCCACTGTGTACAGACAGAGTGGTGTGTACTCTGTGGGATGCATGG 1338
 Qy 341 LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys 360
 Db 1339 CTGNAGACACAGGAATAAGCAATGCTTTGACGCTGCTGGGCAACGAGTGCAGCTGC 1398
 Qy 361 GlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCysValLeu 380
 Db 1399 CAAGAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGAGAGCCTATGTGCTTA 1458
 Qy 381 ProPheThrTyrAsnAspArgThr----- 388
 Db 1459 CCATTTCCTACAAATGGCAGGACGTTCTACTCTCTGCCACCAAGAGGGCGACAGCGGA 1518
 Qy 389 -----AspSerThrThrSerAsnTyrGluGlnAspGlnLysTyrSerPheCysThr 405
 Db 1519 CATCTTTGGTGCACCACTTCGAATATGACAGGACCAAGAAATACTCTTCTGCACA 1578
 Qy 406 AspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGlyAlaLeuCysHisPhe 425
 Db 1579 GACCACACTGTTTTGGTTTCAGACTCGAGGAGGAAATCCAAATGCTGCTTGTGCCACTTC 1638
 Qy 426 ProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGluGlyArgAspAsn 445
 Db 1639 CCCTTCTATACAAACCACTTACACTGATTGCACTTCTGAGGGCAGAGAGACAAAC 1698
 Qy 446 MetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLysPheGlyPheCysPro 465
 Db 1699 ATGAAGTGGTGGGACCAACAGAACTATGATGCCCAAGAGTTGGGTTCTGCCCC 1758
 Qy 466 MetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMetTyrArgIleGlyAsp 485
 Db 1759 ATGCTGCCACAGGAGAAATCTGCACAACTCAAGGGGTGATGTATCCGCTATGGAGAT 1818
 Qy 486 GlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThrCysValGlyAsnGly 505
 Db 1819 CAGTGGGATAGCAGCATGATCGGTGTCATGATGAGGTGCACGTGTGTGGGAAATGGT 1878
 Qy 506 ArgGlyGluTyrThrCysIleAlaTyrSerGlnLeuArgAspGlnCysIleValAspAsp 525
 Db 1879 CGTGGGATAGCAGCATGATCGCTTCTCGCAGCTTCGAGATCAGTGCATGTGTGATGAC 1938
 Qy 526 IleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGluGlyHisMetLeuAsn 545

```
Db 1939 ATCACTTACATGTGAACGACACACATTCACAAAGCGTCATGAAGAGGGGCACATGCTGAAC 1998
Qy 546 CysThrCysPheGlyGlnGlyArgGlyArgGlyPheCysAspProValAspGlnCysGln 565
Db 1999 TGTACATGCTTCGGTCAGGGTCGGGGCAGGTGGAGGTGTATCCCGTCGACCAATGCCAG 2058
Qy 566 AspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGluLysTyrValHisGly 585
Db 2059 GATTACAGAGACTGGGACGTTTTATCAAAATTGGAGATTCATGGGAGAGTATGTGCATGGT 2118
Qy 586 ValArgTyrGlnCysTyrCysTyrGlyArgGlyIleGlyGluTrpHisCysGlnProLeu 605
Db 2119 GTCAGATACCACTGCTACTGCTATGGCCGTGGCATTGGGGAGTGGCATTGCCAACCTTTA 2178
Qy 606 GlnThrTyrProSerSerSerGlyProValGluValPheIleThrGluThrProSerGln 625
Db 2179 CAGACCTATCCAAAGCTCAAGTGGTCTCTGTCAAGTATTTATCACTGAGACTCCGAGTCAG 2238
Qy 626 ProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHisIleSerLysTyrIle 645
Db 2239 CCCAACTCCCACCCCATCCAGTGGATGCACACAGCCATCTCACATTTCCCAAGTACATT 2298
Qy 646 LeuArgTrpArgPro 650
Db 2299 CTCAGGTGGAGACCT 2313
```

Search completed: November 11, 2005, 09:36:50
Job time : 2251.76 secs

This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description	
1	1887.4	87.9	8044	4	US-09-566-921-135	Sequence 135, Appl	
2	1826.8	85.1	7803	2	US-08-551-356-1	Sequence 1, Appli	
3	1826.8	85.1	7803	5	PCT-US93-12687-1	Sequence 1, Appli	
4	1743.4	81.2	7679	4	US-09-220-132-38	Sequence 38, Appli	
5	1743.4	81.2	7680	4	US-09-023-655-1289	Sequence 1289, Ap	
6	1743.4	81.2	7680	5	PCT-US95-09819-6	Sequence 6, Appli	
7	1740.2	81.1	7705	1	US-08-259-569-16	Sequence 16, Appl	
8	1740.2	81.1	7705	2	US-08-826-885-16	Sequence 16, Appl	
9	1738.6	81.0	7705	6	5455158-2	Patent No. 5455158	
10	1738.6	81.0	7705	6	5455158-2	Patent No. 5455158	
11	671.2	31.3	986	1	US-07-637-250A-8	Sequence 8, Appli	
12	671.2	31.3	986	1	US-08-145-061-8	Sequence 8, Appli	
13	97.4	4.5	186	1	US-08-153-799-5	Sequence 5, Appli	
14	73.8	3.4	2109	4	US-09-799-451-345	Sequence 345, App	
15	73.8	3.4	2334	4	US-09-023-655-996	Sequence 996, App	
16	73.8	3.4	2334	4	US-09-949-016-704	Sequence 704, App	
17	73.8	3.4	2335	4	US-09-799-451-346	Sequence 346, App	
18	73.8	3.4	2335	4	US-09-949-016-4758	Sequence 4758, App	
19	73.2	3.4	11665	4	US-09-949-016-12446	Sequence 12446, A	
20	73.2	3.4	11665	4	US-09-949-016-16500	Sequence 16500, A	
21	72.4	3.4	85	1	US-08-259-569-28	Sequence 28, Appl	
22	72.4	3.4	85	2	US-08-826-885-28	Sequence 28, Appl	
23	72.2	3.4	2334	1	US-08-457-304A-33	Sequence 33, Appl	
24	72.2	3.4	2334	1	US-08-456-701A-33	Sequence 33, Appl	
25	72.2	3.4	2334	3	US-08-684-932A-33	Sequence 33, Appl	
26	72	3.4	72	2	US-08-259-569-29	Sequence 29, Appl	
27	72	3.4	72	2	US-08-826-885-29	Sequence 29, Appl	

QY 361 CTGGAAACACTTTACCGAGTGGGTGACACTTATGACGCTCCTAAGACTCCATGATCTGGG 420
Db 573 CTGGGAAACACTTTACCGAGTGGGTGACACTTATGACGCTCCTAAGACTCCATGATCTGGG 632
QY 421 ACTGTACCTGCATCGGGGCTGGCGAGGAGAGAAATAGCTGTACCACTGCAACCGCTGCC 480
Db 633 ACTGTACCTGCATTTGGGGCTGGCGAGGAGAAATAGCTGTACCACTCGCAACCGCTGCC 692
QY 481 ATGAAGGGGGTCACTCTTACAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGGTG 540
Db 693 ATGAAGGGGGTCACTCTTACAAGATTGGTGACACCTGGAGGAGACCAATGAGACTGGTG 752
QY 541 GTTACATGTTAGAGTGTGTCTTGGTAAATGGAAGAGGAAATGGACCTGCAAGCCCA 600
Db 753 GTTACATGTTAGAGTGTGTCTTGGTAAATGGAAGAGGAAATGGACCTGCAAGCCCA 812
QY 601 TAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTTATGTGTGCGGAGAAACGTGGG 660
Db 813 TAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCTTATGTGTGCGGAGAAACGTGGG 872
QY 661 AGAAGCCCTACCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Db 873 AGAAGCCCTACCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 932
QY 721 GCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTATAGAA 780
Db 933 GCATCATTGCACTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTATAGAA 992
QY 781 TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCAGTGCATCTGCACAG 840
Db 993 TTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCAGTGCATCTGCACAG 1052
QY 841 GCAACGGCGGAGAGAGTGAAGTGTGAGAGGACACCTCTGTGACAGCACATCGAGCG 900
Db 1053 GCAACGGCGGAGAGAGTGAAGTGTGAGAGGACACACCTCTGTGACAGCACATCGAGCG 1112
QY 901 GATCTGGCCCTTACCGAGTGTCTGTCAGCTGTTTACCAACCGCAGCTCACCCCCAGC 960
Db 1113 GATCTGGCCCTTACCGAGTGTCTGTCAGCTGTTTACCAACCGCAGCTCACCCCCAGC 1172
QY 961 CTCCTCCCTATGGCCACTGTGTCAAGACAGTGTGTCTACTCTCTGTGGGATGCACT 1020
Db 1173 CTCCTCCCTATGGCCACTGTGTCAAGACAGTGTGTCTACTCTCTGTGGGATGCACT 1232
QY 1021 GGCTGAAGACACAGGAAATAGCAAAATGCTTTTGACAGTGTCTGGGCAACGGAGTCAGCT 1080
Db 1233 GGCTGAAGACACAGGAAATAGCAAAATGCTTTTGACAGTGTCTGGGCAACGGAGTCAGCT 1292
QY 1081 GCCAAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGGCCATGTGTCT 1140
Db 1293 GCCAAGACAGCTGTAAACCCAGACTTACGGTGGCAACTCAAAATGGAGGCCATGTGTCT 1352
QY 1141 TACCATTCACTTACCAACGACAGGAC----- 1165
Db 1353 TACCATTCACTTACCAATGGCAGGACGTTCTACTCTCTGACACACAGAGGGGACAGGAGC 1412
QY 1166 -----GGACAGCAACCTTCGAATTTATGACGAGGACACAGAAATACTCTTTCTGCA 1215
Db 1413 GACATCTTTGGTGACAGCAACCTTCGAATTTATGACGAGGACACAGAAATACTCTTTCTGCA 1472
QY 1216 CAGACCACTGTTTGGTTGACACTTCAGAGGAGGAAATCCAAATGGTGCCTTGTGCCACT 1275
Db 1473 CAGACCACTGTTTGGTTGACACTTCAGAGGAGGAAATCCAAATGGTGCCTTGTGCCACT 1532
QY 1276 TCCCTCTCTTACCAACCAACCACTTACACTGATTGCACTTCTCAGGGGACAGAGAGACA 1335
Db 1533 TCCCTCTCTTACCAACCAACCACTTACACTGATTGCACTTCTCAGGGGACAGAGAGACA 1592
QY 1336 ACATGAAGTGTGTGGGACCAACAGAACTATGATGCCGACACAGAAATTTGGGTTCTGCC 1395
Db 1593 ACATGAAGTGTGTGGGACCAACAGAACTATGATGCCGACACAGAAATTTGGGTTCTGCC 1652

QY 1396 CCATGGCTGCCACGAGGAAATCTGCAACCAATGAAGGGGTCACTGACCGCATTTGGAG 1455
Db 1653 CCATGGCTGCCACGAGGAAATCTGCAACCAATGAAGGGGTCACTGACCGCATTTGGAG 1712
QY 1456 ATCAGTGGGATTAAGCAGCATGACATGGGTCAATGATGAGTGCACGTGTGTTGGGAATG 1515
Db 1713 ATCAGTGGGATTAAGCAGCATGACATGGGTCAATGATGAGTGCACGTGTGTTGGGAATG 1772
QY 1516 GTCTGGGGAAATGGACATGATTCCTTACTCGCAGCTTCGAGATCACTGCAATGTTGATG 1575
Db 1773 GTCTGGGGAAATGGACATGATTCCTTACTCGCAGCTTCGAGATCACTGCAATGTTGATG 1832
QY 1576 ACATCACTTAAATGTGAACGACACACATTCACAAAGCGCTCATGAAGAGGGGCACATGCTGA 1635
Db 1833 ACATCACTTAAATGTGAACGACACACATTCACAAAGCGCTCATGAAGAGGGGCACATGCTGA 1892
QY 1636 ACTGTACATGCTTCGGTTCAGGGTGGGGCAGGTGGAAGTGTGATCCCGTGCACCAATGCC 1695
Db 1893 ACTGTACATGCTTCGGTTCAGGGTGGGGCAGGTGGAAGTGTGATCCCGTGCACCAATGCC 1952
QY 1696 AGGATTCAGAGACTGGGACGCTTTTATCAAAATGGAGATTCATGGGAGAGTATGTCATG 1755
Db 1953 AGGATTCAGAGACTGGGACGCTTTTATCAAAATGGAGATTCATGGGAGAGTATGTCATG 2012
QY 1756 GTGTCAATACCAAGTGTCTACTGCTATGCGCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 1815
Db 2013 GTGTCAATACCAAGTGTCTACTGCTATGCGCGTGGCATTTGGGAGTGGCATTTGCCAACCTT 2072
QY 1816 TACAGACTTATCCAAAGTCAAGTGTGCTCTGTGGAAGTATTTATCACTGAGACTCCGAGTC 1875
Db 2073 TACAGACTTATCCAAAGTCAAGTGTGCTCTGTGGAAGTATTTATCACTGAGACTCCGAGTC 2132
QY 1876 AGCCCAACTCCCAACCCCATCCAGTGGAAATGCACACAGCCATCTCAGATTTCCAAGTACA 1935
Db 2133 AGCCCAACTCCCAACCCCATCCAGTGGAAATGCACACAGCCATCTCAGATTTCCAAGTACA 2192
QY 1936 TTCTCAGTGGAGACCT 1952
Db 2193 TTCTCAGTGGAGACCT 2209

RESULT 2

US-08-551-356-1
; Sequence 1, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Iranl, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A
; REGISTRATION NUMBER: 33-246

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 11, 2005, 08:45:16 ; Search time 17670.7 Seconds
(without alignments)
1809.796 Million cell updates/sec

Title: US-09-581-651D-1

Perfect score: 3781

Sequence: 1 NLVATCLPVRASLPHRLNML.....ISKYILWRPVSIPPRNLGY 660

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_h/US09581651/runat_07112005_092223_28778/app_query.fasta_1.1834
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -WATRIK=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09581651@cgn 1.1 9936 @runat_07112005_092223_28778 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3781	100.0	2147	6 BD137021	BD137021 Polypepti
2	3781	100.0	2147	6 AX003229	AX003229 Sequence
3	3781	100.0	2147	9 HSA276395	AU276395 Homo sapi
4	3774	99.8	2358	6 CQ875358	CQ875358 Sequence

ALIGNMENTS

RESULT 1
BD137021
LOCUS BD137021 2147 bp DNA linear PAT 18-SEP-2002
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION BD137021
VERSION BD137021.1 GI:23231966
KEYWORDS JP 2002508179-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2147)
AUTHORS Schor, S.L. and Schor, A.M.
TITLE Polypeptides, polynucleotides and uses thereof
JOURNAL Patent: JP 2002508179-A 1 19-MAR-2002;
UNIVERSITY OF DUNDEE
COMMENT OS Homo sapiens (human)
PN JP 2002508179-A/1
PF 15-DEC-1998 JP 2000539133
PR 16-DEC-1997 GB 9726539.1
PI SETH LAWRENCE SCHOR, ANA MARIA SCHOR
PC C12N15/09,A61K38/00,A61P17/02,C07K14/78,C07K16/18,C12N5/10, PC
C12P21/08,
PC C12Q1/68,G01N33/574,C12N15/00,A61K37/02,C12N5/00 CC
Polypeptides, polynucleotides and uses thereof. FH Key
Location/Qualifiers 1. .2147
FT source

AJ535086 Homo sapi
BX538045 Homo sapi
AB191261 Homo sapi
CQ833991 Sequence
CQ875357 Sequence
AL833271 Homo sapi
BX64608 Homo sapi
BX640802 Homo sapi
BX649182 Homo sapi
BX454662 Sequence
AX281712 Sequence
BX640875 Homo sapi
BX538017 Homo sapi
BX538018 Homo sapi
BX640731 Homo sapi
CQ871810 Sequence
CQ871828 Sequence
AL833202 Homo sapi
BX537590 Homo sapi
BX640803 Homo sapi
AR051657 Sequence
AR274901 Sequence
AR380744 Sequence
AX277596 Sequence
AX333368 Sequence
X02761 Human mRNA
A14133 Fibronectin
E01162 cDNA encodi
E01162 cDNA encodi
I70110 Sequence 16
AR364992 Sequence
BC051082 Mus muscu
AX402055 Sequence
X15906 Rat mRNA fo
CQ715726 Sequence
CQ72841 Xenopus 1
M77820 Xenopus lae
BV178397 sqm10026
AF081127 Danio rer
AF081128 Danio rer
AY538257 Danio rer

5 3742.5 99.0 2192 9 HSA535086
6 3732.5 98.7 4321 9 HSM806214
7 3693.5 97.7 7753 9 AB191261
8 3693.5 97.7 8037 6 CQ833991
9 3693.5 97.7 8027 6 CQ875357
10 3686.5 97.5 7502 9 HSM804082
11 3686.5 97.5 7868 9 HSM806653
12 3686.5 97.5 7951 9 HSM806901
13 3686.5 97.5 7951 9 HSM806902
14 3686.5 97.5 8044 6 AR454662
15 3686.5 97.5 8044 6 AX281712
16 3686.5 97.5 8411 9 HSM806992
17 3683.5 97.4 8030 9 HSM806170
18 3683.5 97.4 8035 9 HSM806171
19 3682.5 97.4 8042 9 HSM806805
20 3680 97.3 1929 6 CQ871810
21 3680 97.3 1929 6 CQ871828
22 3679.5 97.3 7544 9 HSM803509
23 3674.5 97.2 8320 9 HSM806267
24 3674.5 97.2 8421 9 HSM806903
25 3584.5 94.8 7803 6 AR051657
26 3448.5 91.2 7679 6 AR274901
27 3448.5 91.2 7680 6 AR380744
28 3448.5 91.2 7680 6 AX277596
29 3448.5 91.2 7680 6 AX333368
30 3448.5 91.2 7680 9 HSF1B1
31 3433.5 90.8 7705 6 A14133
32 3433.5 90.8 7705 6 E01162
33 3433.5 90.8 7705 6 E01162
34 3433.5 90.8 7705 6 I70110
35 3428.5 90.7 7705 6 AR364992
36 3423.5 90.5 3059 10 BC051082
37 3381.5 89.4 8329 6 AX402055
38 3381.5 89.4 8329 10 RNFBIRON
39 3389.5 87.0 6816 6 CQ715726
40 2888 76.4 8313 5 BC072841
41 2878 76.1 8216 5 XELFBRNCT
42 2552.5 67.5 7323 11 BV178397
43 2516 66.5 3145 5 AF081127
44 2516 66.5 8267 5 AF081128
45 2332 61.7 8254 5 AY538257

FEATURES		FT	Location/Qualifiers		/organism='Homo sapiens (human)'	
Source		1. .2147	1. .2147		1. .2147	
			/organism="Homo sapiens"		/organism="Homo sapiens"	
			/mol_type="genomic DNA"		/mol_type="genomic DNA"	
			/db_xref="taxon:9606"		/db_xref="taxon:9606"	
ORIGIN						
Alignment Scores:			1.1e-263		Length: 2147	
Pred. No.:			3781.00		Matches: 660	
Score:			100.00%		Conservative: 0	
Percent Similarity:			100.00%		Mismatch: 0	
Best Local Similarity:			100.00%		Indels: 0	
Query Match:			100.00%		Gaps: 0	
DB:			6			
US-09-581-651D-1 (1-660) x BD137021 (1-2147)						
Qy	1	AsnLeuValAlaThrCysLeuProValArgAlaSerLeuProHisArgLeuAsnMetLeu	20			
Db	3	AACITGGTGGGAACTTGCCTCCCGTGGGGGCTCTCTCCCCACCGTCTCAACATGCTT	62			
Qy	21	ArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAlaValPro	40			
Db	63	AGGGTCCGGGCGCGGCTGCTGCTGCGCGCTCCAGTCCCTGGGGACAGCGGTGCC	122			
Qy	41	SerThrGlyAlaSerLysSerLysArgGlnAlaGlnGlnMetValGlnProGlnSerPro	60			
Db	123	TCCACGGGAGCTCGAAGAGGAGGAGGAGGCTCAGCAAAATGGTTCAGCCCACTGCCCG	182			
Qy	61	ValAlaValSerGlnSerLysProGlyCysThrAspAsnGlyLysHisThrGlnIleAsn	80			
Db	183	GTGCTGTCACTCAAGCAAGCCCGTGTGTATGACAAATGGAAACACTATCAGATAAT	242			
Qy	81	GlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysThrGlyCysSer	100			
Db	243	CAACAGTGGGAGCGGACCTACTAGGCAATGCGTTGGTTGTACTTTATGGAGGAGC	302			
Qy	101	ArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLysThrThr	120			
Db	303	CGAGGTTTAACTGCGAGAGTAAACCTGAAGCTGAAGAGACTTGCCTTTGACAAGTACACT	362			
Qy	121	GlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIleTrpAsp	140			
Db	363	GGGAACACTTACCGAGTGGTGACACTTATGAGCGTCTTAAAGACTCCATGATCTGGAC	422			
Qy	141	CysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArgCysHis	160			
Db	423	TGTACCTGCATCGGGCTGGGCGAGGAGAGATAAGCTGTACCATCGCAACCGCTGCCAT	482			
Qy	161	GluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThrGlyGly	180			
Db	483	GAAGGGGTGAGTCTCTACAAGATTGGTGACACCTGGAGGAGACCATGAGACTGGTGGT	542			
Qy	181	TyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLysProIle	200			
Db	543	TACATGTTAGAGTGTGTGTCTTGGTATGGAAGGAGAGATGAGCTGCAAGCCCAITA	602			
Qy	201	AlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGlu	220			
Db	603	GCTGAGAAGTGTGTTTGTATCATGCTCTGGGACTTCTTATGTTGGTGGGAGAAACGTGGGAG	662			
Qy	221	LysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArg	240			
Db	663	AAGCCCTACCAAGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	722			
Qy	241	IleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIle	260			
Db	723	ATCATTCTGCATCTTAGAATAATAGATGCAACGATGAGGACACAGAGACATCTCTATAGATT	782			
Qy	261	GlyAspThrTrpSerLysAspAsnArgGlyAsnLeuLeuGlnCysIleCysThrGly	280			
Db	783	GGAGACACTGGAGCAAGAGGATTAATCGAGGAAACCTGCTCCAGTGCATCTGCACAGGC	842			
Qy	281	AsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGly	300			
Db	843	AACGGCCGAGAGAGTGGAAAGTGTGAGAGCAGACACCTCTGTGTGAGACCATCGAGCGGA	902			
Qy	301	SerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisProGlnPro	320			
Db	903	TCTGGCCCTTACCCGATGTTCTGTCAGCTGTTTACCAACCGCAGCCTCACCCCGCCT	962			
Qy	321	ProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMetGlnTrp	340			
Db	963	CCTCCCTATGGCCACTGTGTACAGACAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1022			
Qy	341	LeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyValSerCys	360			
Db	1023	CTGAAGACACAGGAAATAAGCAATGCTTTTGGACGTGCTGGGCAACGAGTCTGCTGC	1082			
Qy	361	GlnGluThrAlaValThrGlnThrTyrGlyLysAsnSerAsnGlyGluProCysValLeu	380			
Db	1083	CAAGAGACAGCTGTAACCCAGACTTACGGTGGCAACTCAAAATGGAGAGCCATGTGTCTTA	1142			
Qy	381	ProPheThrTyrAsnAspArgThrAspSerThrThrSerAsnTyrGluGlnAspGlnLys	400			
Db	1143	CCATTTCACCTACAACGACGAGCAGCAGCAACTTCGAAATTTATGACGAGGACCGAGAA	1202			
Qy	401	TyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSerAsnGly	420			
Db	1203	TACTCTTTTCGACAGACCACTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	1262			
Qy	421	AlaLeuCysHisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThrSerGlu	440			
Db	1263	GCCTTGTGCCACTTCCCTTCTTATACAACAACCAAAATTTACACTGATTTGCTCTGAG	1322			
Qy	441	GlyArgArgAspAsnMetLysTrpCysGlyThrThrGlnAsnTyrAspAlaAspGlnLys	460			
Db	1323	GGCAGAGAGACAACTGAAGTGGTGGGACCAACAGCACTATGATGGCGACCGAGAG	1382			
Qy	461	PheGlyPheCysProMetAlaAlaHisGluGluIleCysThrThrAsnGluGlyValMet	480			
Db	1383	TTTGGGTTCTGCCCTTGGCTGCCACAGAGAAATCTGCACACCAATGAAGGGGTCTATG	1442			
Qy	481	TyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArgCysThr	500			
Db	1443	TACCGCATTTGAGATCATGATGGGATAAGCAGCATGACATGGGTGCATGATGAGGTGCACG	1502			
Qy	501	CysValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArgAspGln	520			
Db	1503	TGTGTTGGGAATGGTTCGGGGAATGGACATGTCATTCCTCTACTCGCAGCTTCGAGATCAG	1562			
Qy	521	CysIleValAspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHisGluGlu	540			
Db	1563	TGCATTGTTGATGACATCATTTACATGTAACAGACACATTCACACAGCGTCTGAGAGAG	1622			
Qy	541	GlyHisMetLeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCysAspPro	560			
Db	1623	GGGCACATGCTGAACCTGATACATGCTTCCGTCCGGTCCGGGACGGTGGAGTGTGATCCC	1682			
Qy	561	ValAspGlnCysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSerTrpGlu	580			
Db	1683	GTCACCAATGCCAGGATTCAGAGACTGGGACGTTTATCAAAATGGAGATTCATGGGAG	1742			
Qy	581	LysTyrValHisGlyValArgTyrGlnCysThrCysThrGlyArgGlyIleGlyGluTrp	600			
Db	1743	AAGTATGTGATGTTGTCAGATACAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	1802			
Qy	601	HisCysGlnProLeuGlnThrTyrProSerSerSerGlyProValGluValPheIleThr	620			
Db	1803	CATTGCCAACCTTTACAGACCTATCCAAGCTCAAGTGTCTGTCTGTCTGTCTGTCTGTCTGT	1862			
Qy	621	GluThrProSerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnProSerHis	640			
Db	1863	GAGACTCCGAGTCCAGCCCACTCCACCCCATCCAGTGGAAATGACACACACCATCTCTCAC	1922			